



K1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention
of the grant of the patent:
04.09.1996 *Bulletin* 1996/36

(51) Int Cl.⁶ **C12Q 1/70, C12Q 1/68,**
C12P 19/34
// C07H21/04

(21) Application number **91116561.1**

(22) Date of filing: **27.09.1991**

(54) **Nucleotide sequences useful as type-specific probes, PCR primers and LCR probes for the amplification and detection of human papilloma virus, and related kits and methods**

Nukleotid-Sequenzen nützlich als typenspezifische Sonden, PCR Primers und LCR Sonden zur Amplifikation und zum Nachweis von humanem Papillomavirus, sowie dazu verwendete Kits und Verfahren

Séquences nucléotidiques utiles comme sondes spécifiques du type amorces de PCR et sondes pour l'amplification et détection du virus-papilloma humain, et kits et procédés utilisés dans ce but

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IT LI NL SE

(30) Priority: **28.09.1990 US 589948**
28.09.1990 US 590105
28.09.1990 US 590253

(43) Date of publication of application:
01.04.1992 *Bulletin* 1992/14

(73) Proprietor: **ABBOTT LABORATORIES**
Abbott Park, Illinois 60064-3500 (US)

(72) Inventors:
• **Joseph, Jeffrey L.**
Cherry Hill, New Jersey 08002 (US)
• **Bouma, Stanley R.**
Mundelein, Illinois 60060 (US)

• **Marshall, Ronald L.**
Zion, Illinois 60099 (US)
• **Laffier, Thomas G.**
Libertyville, Illinois 60048 (US)

(74) Representative: **Modiano, Guido, Dr.-Ing. et al**
Modiano & Associati S.r.l.
Via Meravigli, 16
20123 Milano (IT)

(56) References cited:
EP-A-0 301 968 **EP-A-0 320 308**
EP-A-0 402 132 **EP-A-0 425 995**
WO-A-87/05630 **WO-A-89/05357**
WO-A-89/09940 **WO-A-90/02821**
WO-A-91/10675

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention)

Description

This invention relates generally to human papilloma virus, and more particularly, relates to nucleotide sequences of short strands of human papilloma virus which can be amplified and/or used to determine the presence of human papilloma virus products in a test sample, and some of which also can be amplified and/or used to determine the specific type of human papilloma virus of types 16 and 18 present in the test sample.

Human papilloma virus (HPV) is recognized as a venereally-transmitted disease of the anogenital tract which often is associated with the pathogenesis of cervical cancer and its precursor lesions. More than 55 types of HPV have been characterized. Of these, at least 21 types infect the anogenital tract. L. Gregoire et al., *J. Clin. Microb.* 27 (12) 2660-2655 (1989). These mucosotropic viruses are associated most frequently with benign condyloma or latent infections. However, the presence of HPV in premalignant lesions and invasive cancers, particularly of the cervix, may reflect the oncogenic potential of these viruses. See P. M. Howley, in *Important Advances in Oncology*, D. T. DeVita, Jr. et al., eds., J. B. Lippincott, Philadelphia, PA (1987) at pages 55-73.

Certain HPV types, namely, HPV type 16 and type 18, and to a lesser extent HPV types 31, 33 and 35, are found in a high proportion of invasive cervical cancers and their metastases. However, many HPV types which infect the anogenital tract, such as HPV types 6 and 11, are found most commonly in benign condyloma and only rarely are found in invasive cancers. HPV detected in the anogenital tract can be classified broadly as low risk papilloma viruses (HPV types 6 and 11), intermediate risk papilloma viruses (HPV types 31, 33 and 35) or high risk papilloma viruses (HPV types 16 and 18), based on the association of the particular HPV type with malignancy. A. T. Lorincz et al., *J. Nat'l Cancer Inst.* 79 671 (1987). Thus, the detection of the presence of HPV and the determination of the specific type of HPV can provide a diagnostic and prognostic tool useful for determining the clinical significance associated with certain HPV types. The early detection of HPV by sensitive and specific reagents and methodologies also could provide earlier therapeutic management and counseling.

A need therefore exists for accurate and reliable methods to identify and type HPV in clinical specimens. However, known polyclonal antisera prepared by immunizing animals with disrupted virions are capable of detecting HPV antigens in only about 30-70% of cutaneous and mucosal warts. Further, the antisera are broadly cross-reactive. Available immunological tests have two major drawbacks. First, only well-differentiated cells apparently are capable of viral antigen expression. HPV-infected tissues which show higher degrees of neoplasia, such as carcinoma *in situ*, rarely contain HPV antigen. Thus, the further the development of the malignancy, the smaller the amount of detectable virus in the tested tissue. Secondly, these immunological tests are unable to identify specific viral types.

It is known that papilloma viruses share amino acid sequences in the major capsid proteins. See, for example, C. C. Baker, in *The Papovaviridae* (Vol. 2), P. M. Howley and N. P. Salzman, eds., Plenum Publ. Corp., New York (1987) at pages 321-385. The DNAs of this virus cross-hybridize, indicating homologous sequences. M. F. Law et al., *J. Virol.* 58 225-229 (1979). Thus, molecular hybridization techniques have been developed as a more sensitive and specific means of detecting and differentiating HPV DNA and RNA in clinical specimens. See A. T. Lorincz, *Gynecol. Clinics of N. America* 14 451 (1987).

Sequences specific for the DNA and RNA of human papilloma virus are known and have been published. See, for example, PCT application No. WO 89/69940 published October 19, 1989, PCT application No. WO 89/05216 published October 9, 1989 and European Patent Application No. 0 301 968 published February 1, 1989.

The molecular hybridization techniques used to detect homologous DNA sequences are sensitive and can be highly specific if used with probes which bind to nucleic acid sequences which are unique to a particular HPV type. However, the concentration of total viral DNA in a given clinical sample may be below the limit of sensitivity of the test. For example, the amount of viral DNA in dysplastic cervical lesions is reduced with increasing dysplasia.

To overcome this problem of sensitivity, viral DNA sequences can be amplified by using, for example, the polymerase chain reaction (PCR) or the ligase chain reaction (LCR) techniques. The products thus obtained can be identified by using conventional hybridization techniques for identification of virus types, such as Southern blotting. See C. Ostle, *Biotechniques* 6:163 (1988); K. B. Mullis, U. S. Patent No. 4,683,202, and EP-A-320 308 (BioTechnica).

Both PCR and LCR serve to amplify the DNA present in a test sample to detectable levels. In practice, the level of sensitivity is about 50 to 100 copies per sample. The next most sensitive technique is dot-blot, which can detect about 10,000 molecules, while Southern blot reliably detects about 100,000 copies of DNA per sample.

Thus the appropriate diagnosis of HPV may require two steps. In one strategy, the presence of a clinically relevant type of HPV is first detected with a group-specific primer. After the presence of HPV is detected, differentiation between types can be performed by using a type-specific probe having low homology between the HPVs of the group. Alternatively, differentiation can be performed using a mixture of type-specific probes at the outset, provided these probes amplify DNA independently of each other, and that they can be detected independently. In the past, such tasks were attempted using specific antibodies. In general, nucleic acid probes and primers allow greater discrimination among subtypes than do antibodies. The use of DNA-based tests increases both sensitivity and specificity over prior-art antibody-based tests.

It therefore would be advantageous to provide oligonucleotide strands of DNA which could be amplified and used to detect the presence, if any, of HPV in a test sample. It also would be advantageous to provide short oligonucleotide strands of DNA which could be amplified and used to detect the presence, if any, of specific types of HPV in the test sample. The combined use of oligonucleotide strands would be advantageous for allowing for the specific and sensitive *in vitro* diagnosis of the presence and specific type of HPV present in test samples.

SUMMARY OF THE INVENTION

Oligonucleotides of from about 10 to about 60 nucleotides are provided which can be amplified and used either to detect specific sequences of specific types of human papilloma virus, or consensus regions with high homology among different types. The presence of HPV is determined by contacting the test sample with sequences provided to detect the presence, if any, of HPV types 6, 11, 16, 18, 31, 33 and 61. This may be done with or without prior amplification, for example, by PCR or LCR. Either type-specific or consensus amplification is also possible. Two oligonucleotides are provided if the sequence is to be amplified by PCR, and four oligonucleotides provided if amplification is by LCR, in accordance with these known amplification procedures. After the presence of HPV is detected, the type of HPV present in the sample can be determined by using HPV type-specific probes, by subsequent rounds of PCR, or by LCR. Alternatively, the presence of type-specific HPV can be determined by contacting the test sample directly with type-specific nucleotide sequence provided by the invention for the detection of HPV types 16 and 18. Also provided are methods for using the oligonucleotides and kits for amplifying and detecting the presence of human papilloma virus.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a photograph of a gel following electrophoresis showing the results when the primers PCR 1 and PCR5 were used to amplify selected plasmids wherein HPV 6 is in lane 1, HPV 11 is in lane 2, HPV 16 is in lane 3, HPV 18 is in lane 4, and HPV 31 is in lane 5. HPV 33 is in lane 6, HPV 61 is in lane 7, and molecular weight standards are in lane 8.

FIG. 2 is a photograph of a gel following electrophoresis showing the results when the primers PCR 1, PCR2, PCR3, PCR4 and PCR5 were used to amplify plasmid p65 16.8 (HPV 16). PCR1 and PCR5 are primers according to the invention.

FIG. 3 is a photograph of the ethidium bromide-stained gels wherein PCR 1 4 and PCR15 are used in conjunction with IWDO to obtain amplified PCR product.

FIG. 4 is a graph of results obtained from performing LCR on 10^7 molecules of the selected target using LCR6A, LCR5A', LCR6B and LCR5B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

FIG. 5 is a graph of results obtained from performing LCR on 10^7 molecules of the selected target using LCR6A, LCR6A', LCR6B and LCR6B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

FIG. 6 is a graph of results obtained from performing LCR on 10^7 molecules of the selected target using LCR7A, LCR7A', LCR7B and LCR7B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

FIG. 7 is a graph of results obtained from performing LCR on 10^7 molecules of the selected target using LCR8A, LCR8A', LCR8B and LCR8B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

DETAILED DESCRIPTION OF THE INVENTION

The appropriate diagnosis of HPV requires two sets of conditions. The first enables the detection of all pertinent types, and the second set allows differentiation among them. In the past, such tasks have been attempted using specific antibodies. In general, nucleic acid probes and primers allow greater discrimination among subtypes than do antibodies. Thus, the use of DNA-based tests tends to increase both sensitivity and specificity over antibody-based tests.

U. S. Patents No. 4,683,195 and 4,683,202 teach a method of amplifying DNA sequences by using PCR. This method now is a standard procedure in many molecular biology laboratories. Examples 1-3 which follow below utilize the procedures taught in these two patents and the method as described in the package insert of the commercially-available Gene-Amp™ kit (Document No. 55635-5/89, Perkin-Elmer/Cetus, Emeryville, CA).

In PCR, two complementary polynucleotide strands are amplified by treating the strands with two oligonucleotide primers such that an extension product of each primer is synthesized which is complementary to each nucleic acid strand. The primers are selected such that the extension product of one primer forms a template for the synthesis of an extension product from the other primer once the extension product of the one primer is separated from the template. A chain reaction is maintained by a cycle of denaturing the primer extension products from their templates, treating

the single-stranded molecule generated with the same primers to re-anneal, and allowing the primers to form further extension products. The cycle is repeated for any many times as it takes to increase the target nucleic acid segments to a concentration where they can be detected.

The amplified target sequence can be detected by any of several known techniques, for example, by denaturing the double-stranded products formed by PCR, and treating those products with one or more reporter probes which hybridize with the extension products. The reporter probe has a detectable label, and usually is added in excess. The unhybridized reporter probe, therefore, must be separated from the hybridized reporter probe by involving a separation step. In another method of detecting the extension products without reporter probe and a separation step, the extension products are detected by gels stained with ethidium bromide. The diagnosis can be confirmed by transferring the DNA to nitrocellulose and probing with a probe specific to the HPV type suspected of being present in the sample.

Alternately with PCR, one may take advantage of known restriction sites within the HPV DNA to demonstrate that the amplified DNA contains the expected sequence by examining the cleavage pattern(s) generated with one or more restriction endonucleases. Verifying the authenticity of the amplified sequence may be necessary for two reasons: (1) to ensure that sequences complementary to the amplifying primers are not fortuitously present in cellular DNA which does not contain HPV DNA, and (2), to identify the type of HPV present in the sample. If the sequences chosen for amplification are conserved among HPV types, then the finding of an amplified product does not implicate a particular HPV type. It also should be possible to predict the size of the amplified product based on the binding positions of the two primers. Thus, when that product is found, one reasonably can be assured that HPV is present. However, two different types of HPV may give the same or different size products. Thus, hybridization should be used to confirm the identity of the amplified sequence until confidence is built that the interpretation of the results is reliable. It should be pointed out that the PCR technique will identify only closely related, or type-specific sequences in the absence of highly homologous primers, since only a small portion of the genome is analyzed.

Another particularly useful detection technique is described in EP-A-357 011. In this method, a different reporter molecule, e.g. hapten, is attached to each primer. Following amplification, but before denaturation, duplexes can be detected by "capturing" one hapten (hapten1) with a solid phase coated with anti-hapten1. The separated complex can be detected with a conjugate of label and anti-hapten2, and label associated with the solid phase can be measured.

The Ligase Chain Reaction (LCR) amplifies sections of DNA by copying the section of DNA, and copying the copies of that section of DNA, many times over. This method is described in European Patent Application No. 0 320 308 published June 14, 1989, which is incorporated herein by reference. In this procedure, two probes (for example, A and B) complementary to immediately adjacent regions of a target sequence are hybridized and ligated. This ligated probe then is denatured away from the target, after which it is hybridized with two additional probes (A' and B') of sense opposite to the initial probes A and B. The secondary probes are themselves then ligated. Subsequent cycles of denaturation/hybridization/ligation create the formation of double-length probes of both sense (+) and antisense (-).

In LCR, the nucleic acid of the sample is provided either as single stranded DNA or as double-stranded DNA which is denatured to separate the strands. Four probes are utilized: the first two probes (A and B) are the so-called primary probes, and the second two probes (A' and B') are the so-called secondary probes. The first probe (A) is a single strand capable of hybridizing to a first segment of the primary strand of the target nucleotide sequence. The second probe (B) is capable of hybridizing to a second segment of the primary strand of the target nucleotide sequence. The 5' end of the first segment of the primary strand of the target is positioned relative to the 3' end of the second segment of the primary strand of the target to enable joining of the 3' end of the first probe to the 5' end of the second probe, when the probes are hybridized to the primary strand of the target nucleotide sequence. The third probe (A') is capable of hybridizing to the first probe, and the fourth probe (B') is capable of hybridizing to the second probe (B). The hybridized probes are ligated to form reorganized fused probe sequences. Then, the DNA in the sample is denatured to separate ligated probes from sample DNA. Successive cycles wherein the ligated probes and target DNA undergo the above-described process are performed to increase the amount of detectable DNA in the sample. The amount of cycles performed is dependent upon the sequence used and the sensitivity required of the test. Usually, the cycle can be repeated from 15 to 60 times. At least one of the probes can be conjugated to a signal generating compound.

If the four probes are conjugated to appropriate binding members, the detection of amplified product can be accomplished using standard manual or automated immunoassay procedures known to those skilled in the art. These procedures include, for example, immunochromatography, ELISA, EIA and MEIA. Hybridization also can be accomplished by following standard dot-, slot- or replica-blot procedures which are known to those in the art. The sequences can be labelled with an appropriate signal generating compound (label), which is capable of generating a measurable signal detectable by external means. The various signal generating compounds contemplated include chromogens, catalysts such as enzymes, luminescent compounds such as fluoroscein and rhodamine, chemiluminescent compounds, radioactive elements such as ^{32}P , and other labels known to those of ordinary skill in the art. The selection of a particular label is not critical, but it will be capable of producing a signal either by itself or in conjunction with one or more additional substances. A variety of different indicator reagents can be formed of label and specific binding member. Either the label or a specific binding member can be varied. Examples of specific binding members which

can be used as a member of the indicator reagent include antibodies, both monoclonal, polyclonal, and fragments thereof, avidin or biotin, biotin and anti-biotin, a carbohydrate or a lectin, a complementary nucleotide sequence, an effector or a receptor molecule, an enzyme cofactor or an enzyme, an enzyme inhibitor or an enzyme, also any antigenic substances, haptens, antibodies, and combinations thereof.

The test sample can be any biological material suspected of containing HPV. Thus, the test sample can be human body tissue, or a test sample which contains cells suspected of containing HPV.

The invention will now be described by way of Examples, which are meant to describe, but not to limit, the spirit and scope of the invention.

The following terms used in the examples are trademarks, tradenames or chemical abbreviations as specified

TRIS - chemical abbreviation for [tris(hydroxymethyl)aminomethane], used as a buffer

EDTA - chemical abbreviation for ethylenediaminetetraacetic acid, a chelating agent.

FITC - chemical abbreviation for fluorescein isothiocyanate, a fluorescent hapten derivative

NHS-ester - chemical abbreviation for N-hydroxysuccinamide ester

MES - chemical abbreviation for [2-(N-morpholino)ethanesulfonic acid], a buffer

TWEEN®-20 - trademark of Atlas Chemical for polyoxyethylene sorbitan monolaurate, a detergent.

BIS-TRIS - chemical abbreviation for [bis-(2-hydroxyethyl)-amino]tris-(hydroxymethyl)methane, a buffer

TRITON X-100® - trademark of Rohm & Haas for nonaethylene glycol octylphenol ether, a detergent

IMx® - trademark of Abbott Laboratories for an automated instrument for performing microparticle enzyme immunoassay (MEIA).

EXAMPLES

EXAMPLE 1

PCR was performed essentially following the package insert of the commercially available Gene-Amp™ kit (document No. 55635-6/89, available from Perkin-Elmer/Cetus, Emeryville, CA). The following reagents were mixed in a 0.5 mL polypropylene tube and used in performing PCR

Reagent	Final Concentration
Water	(to give final volume = 50 or 100 µL)
Reaction Buffer	10 mM TRIS pH 8.3 50 mM KCl 1.5 mM MgCl ₂ 0.01% gelatin
dNTP mixture	200 µM each of dATP, dCTP, dGTP, and TTP
pCR1	1 µM
pCR2	1 µM
plasmid	10 µL 1 ng/100µL
(or control-human placental DNA (Pooled Placental DNA, catalog D-3287, Sigma Chemical Co, St. Louis MO).	
DNA polymerase, <i>Thermus Aquaticus</i>	25 or 63.9 units/1 mL

After mixing, the reaction mixture was overlaid with 100 µL of mineral oil. The tube then was placed in an instrument capable of incubation at several temperatures, and subjected to 30 or 40 cycles of programmed temperature change. The precise cycle of temperature change used, and the instrument used, varied with the experiment, and is detailed in the descriptions of the figures in Example 3.

EXAMPLE 2

Following the procedure of Example 1, the following sequences were found to amplify sections of papilloma virus types 6, 11, 16, 18, 31, 33, and 61 using PCR.

PCR1: CAGAIGTCTC TGTGGCGGCC TAGTG (10 Ng 1)

PCR5 AGGTGTCAGG AAAACCAAT TTATT (ID No. 5)

PCR14 GAATTAGTGA GACCATTTAA AAG (ID No. 6)

PCR15. GGGGAAACAC CAGAATGGAT A (ID No. 7)

IWDO ATCATATGCC CACTGTACCA T (ID No. 8)

Sequence IWDO is derived from a sequence disclosed in International application number PCT/US86/00629 (WO 86/05816)

TABLE 1 shows the sequences and where they map to in the various types.

TABLE 1
SEQUENCES WHICH CAN BE USED AS PROBES OR PCR PRIMERS

PROBE ID NO.	SEQ ID NO.	SEQUENCE	SENSE	MAPS TO:	MAPS TO:	MAPS TO:	MAPS TO:	MAPS TO:	MAPS TO:
				(type 6)	(type 11)	(type 16)	(type 18)	(type 31)	(type 33)
PCR1	1	CAGATGTCTGTGTGCGCCCTAGT	+	5785-5810	5768-5792	5634-5658	5610-5634	5550-5574	5591-5615
PCR2	2	CGTTTTCATATTTTITTCAGATG	+	5767-5791	5749-5773	5615-5639	5591-5615	5531-5555	5572-5596
PCR3	3	AACTTGTAAACACCAAGATATGT	+	5844-5868	5826-5850	5695-5719	5671-5695	5611-5635	5652-5676
PCR4	4	AATGTACCTTAATACCTATATG	-	6008-5984	5990-5966	5855-5831	5831-5807	5784-5760	5825-5801
PCR5	5	AGGTGTCAGGAAAACCAATTTATT	-	6044-6020	6026-6002	5901-5877	5877-5853	5820-5796	5861-5837
PCR14	6	GAATTAGTGAACCATTTAAAG	+	1495-1517	1495-1517	1524-1546	1595-1617	1462-1484	1518-1540
PCR15	7	GGGAAACACCAAGATGATA	+	1834-1854	1834-1854	1863-1883	1934-1954	1801-1821	1857-1877
IWDO	8	ATCATATGCCACTGTACCA	-	1931-1911	1931-1911	1960-1940	2031-2011	1898-1878	1954-1934

note: PCR2, PCR3 and PCR4 are not probes or PCR primers of the invention

EXAMPLE 3

Linearized plasmids containing full-length papilloma virus inserts in pGEM3 were used as targets. These were pHPV6.1 (HPV6), pSP65.11.5 (HPV 11), p65.16.8 (HPV16), pHPV18H(HPV18), pG3 HPV31 (HPV31), pLNK322.HPV33 (HPV33), and pBR322 HPV81 (HPV81). The Programmable Cycler Reactor™ (available from Ericomp, San Diego) was used as the incubation instrument. Following PCR procedures as described in Example 1, 10 µL aliquots were analyzed by electrophoresis through agarose (comprising a 5:3 ratio of NuSieve® SeaKem® GTG, available from the FMC Corp., Rockland, ME) in a buffer comprising 0.089 M TRIS, 0.089 M borate, 2 mM EDTA, and 0.5 ppt ethidium bromide.

FIG 1 is a photograph of an ethidium bromide-stained 1.2% agarose gel showing results using 63.9 units/mL DNA polymerase, in the DNA Thermal Cycler™ (Perkin-Elmer/CETUS, Emeryville, CA). The samples were heated for 5 minutes at 94°C, then subjected to 40 cycles of a temperature program of: 1 minute at 94°C, 2 minutes at 40°C, and 1.5 minutes at 72°C. The PCR primers used in this case were PCR 1 and PCR5 of Example 2. Examination of the gel following electrophoresis showed bands at the expected positions, i.e. 292 bp. Lane 1: HPV6, lane 2: HPV 11, lane 3: HPV16, lane 4: HPV 18, lane 5: HPV31, lane 6: HPV33, lane 7: HPV61, lane 8: pooled human placental DNA (suspected of having HPV infection), lane 9: molecular weight markers Hae III digest of ΦX174.

FIG 2 is a photograph of an ethidium bromide-stained 4% agarose gel showing results using 25 units/mL DNA polymerase, in the Programmable Cycler Reactor™ (Ericomp, San Diego, CA). Samples in this case were subjected to 30 cycles of a temperature program of: 50°C for one (1) minute, 72°C for two (2) minutes and 95°C for one (1) minute. In this case, the primers PCR1, PCR2, PCR3, PCR4 and PCR5 of Example 2 were used to amplify plasmid

p65, 16, 8 (HPV 16). Examination of the gel of Figure 2 shows bands at the expected positions, i.e. PCR 1 and PCR4, 235 bp, lane 2; PCR1 and PCR5, 267 bp, lane 4; PCR2 and PCR4, 254 bp, lane 6; PCR2 and PCR5, 286 bp, lane 8; PCR3 and PCR4, 174 bp, lane 10; PCR3 and PCR5, 206 bp, lane 12; molecular weight marker, 123, 246, 369, 492, ... bp ladder, lane 1. Note footnote to Table 1

FIG. 3 is a photograph of an ethidium bromide-stained 1.2% agarose gel showing results using the same conditions as FIG. 1. In this case, PCR14 and PCR15 were used as primers in conjunction with IWDO. The expected size of the amplified PCR product of PCR 14 and IWDO is 437 bp for all of the HPV types tested. The expected size of the product of PCR 15 and IWDO is 98 bp. Products of these sizes appear in the gels, confirming that PCR14 and PCR15, used in conjunction with IWDO, will amplify HPV DNA of types 6, 11, 16, 18, 31, 33, and 61. Lane 1, Molecular weight marker (Hae III digest of FX 174). PCR 14 + IWDO, lanes 2-9: lane 2, HPV6, lane 3, HPV 11, lane 4, HPV16, lane 5, HPV18, lane 6, HPV31, lane 7, HPV33, lane 8, HPV61, lane 9, human placental DNA suspected of being infected with HPV. PCR 5 + IWDO, lanes 10-17: lane 10, HPV6, lane 11, HPV 11, lane 12, HPV16, lane 13, HPV18, lane 14, HPV31, lane 15, HPV33, lane 16, HPV61, lane 17, human placental DNA suspected of being infected with HPV, lane 18, molecular weight marker (Hae III digest of FX174 and Hind III digest of 1 DNA).

EXAMPLE 4

The following reagents were mixed in a 0.5 mL polypropylene tube as follows for the Ligase Chain Reaction (LCR)

Reagent	Volume	Final Concentration
Water	21 μ L	
Reaction Buffer	10 μ L	50 mM EPPS pH7.8 10 mM NH_4Cl 10 mM MgCl_2 100 mM K^+ (from all sources) 0.001% BSA 1 mM DDT
Nicotine Adenine Dinucleotide (NAD)	0.5 μ L	100 μ M
Probe A (sense)	4 μ L	5.0×10^{11} molecules
Probe A' (antisense, 5'-phosphate)	4 μ L	7.5×10^{11} molecules
Probe B (sense, 5'-phosphate)	4 μ L	7.5×10^{11} molecules
Probe B' (antisense)	4 μ L	5.0×10^{11} molecules
Target (including human placental carrier DNA at 10 μ g/mL)	1.5 μ L	15 ng/50 μ L
DNA ligase, <i>Thermus thermophilus</i>	1 μ L	

This reaction mixture was overlaid with 30 μ L of mineral oil. The tube was placed in an instrument capable of incubation at several temperatures (e.g. thermal cycler from Coy Laboratory Products (Ann Arbor, MI) or the Programmable Cycler Reactor™ (available from Ericomp, San Diego, CA), and then subjected to several cycles of programmed temperature change. Each cycle involved incubation at 50°C for one minute and 85°C for one minute.

EXAMPLE 5

The following procedure was used when performing the Ligase Chain Reaction (LCR), which is described in published European Patent Application No. 0 320 308 A2. The reagents of Example 4 were utilized in the procedure as follows. Two probes (A and B) complementary to immediately adjacent to regions of a target sequence were hybridized and ligated. This ligated probe was denatured away from the target, and hybridized with two additional probes (A' and B') of sense opposite to the initial probes (A and B). The secondary probes then were ligated. Subsequent cycles of denaturation/hybridization/ligation created the formation of double-length probes of both + and - sense.

EXAMPLE 6

The following sequences were determined to be specific for a portion of the E6 region of HPV type 16:

Probe	SEQ ID No.	Sequence	Maps to
LCRSA	81	GCTGCAACA ACTATACATG ATATAA	157 - 182
LCRSA'	82	TTATATCATG TATAGTTGTT TGACG	182 - 157
LCRSB	83	TTATTAGAATG TGTGACTGC AAGCA	183 - 208
LCRSB'	84	TGCTTGCACT ACACACATTC TAATA	208 - 157

EXAMPLE 9

Base-denatured plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. These plasmids were pG3HPV6(+) (HPV6), pSP 65, 11.5 (HPV11), pSP65 168 (HPV16), pG3HPV18(-)(HPV18), p63 HPV31 (HPV31), pLNK322 HPV33 (HPV33), pBR322 HPV35 (HPV35), pUC19 HPV52 (HPV52), pLNK322 HPV58 (HPV58), pUC9 HPV59 (HPV59) and pBR322 HPV61 (HPV61). All of the oligonucleotides used as probes from Example 8 had chemical labels covalently attached at the ends distal from ligation. These labels were: 5'-fluorescein-LCRSA, 3'-fluorescein-LCRSA', 3'-biotin-LCRSB and 5'-biotin-LCRSB'. Covalent attachment was performed by known methods, i.e., reaction of amine-terminated oligonucleotides with FITC or biotin-NHS-ester essentially following the procedures of Kansal et al. Tel. Letters 29 5537-5540 (1988). The thermal cycler used was obtained from Coy Laboratory Products, Ann Arbor, MI.

Following the LCR procedure of Examples 4 and 5, the mixtures were analyzed using a prototype version of the IM₂ instrument (Abbott Laboratories, Abbott Park, IL), following the protocol for microparticle enzyme immunoassays as follows: A 40 µL aliquot of an LCR mixture was diluted 1:1 with distilled water. This diluted mixture was incubated with 50 µL anti-fluorescein-conjugated polystyrene microparticles for five (5) minutes to form a suspension of immune complexes on the microparticles. This suspension then was transferred to an inert glass fiber matrix, to which the microparticles became attached. The matrix was washed with buffer (0.3M NaCl, 10 mM TRIS pH8, 0.1%Na₂S₂O₅). Any immune complexes attached to the glass matrix was detected by using alkaline phosphatase-labeled conjugate that catalyzed the hydrolysis of 4-methylumbelliferone. The rate at which the 4-methylumbelliferone was generated on the matrix was proportional to the concentration of LCR product formed in the reaction mixture.

Referring to FIG. 4, the graph shows the results obtained from performing LCR on 10⁷ molecules of the targets in shown. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second/second. Background signal is approximately 10 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV16, and those values are about 60 times background signal.

EXAMPLE 10

The following sequences were determined to be specific for a portion of the E6 region of HPV type 18:

Probe	SEQ ID No.	Sequence	Maps to
LCR6A	85	CTTCACTGCA AGACATACAA ATAA	172 - 195
LCR6A'	86	TTATTTCTAT GTCTTGCACT GAA	195 - 173
LCR6B	87	CCCTGTGTATA TTGCAAGACA GTAT	196 - 219
LCR6B'	88	TACTGTCTTG CAATATACAC AGG	218 - 196

EXAMPLE 11

Plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. The plasmids used were those described in Example 9. All of the oligonucleotides used as probes obtained from Example 10 had chemical labels covalently attached at the ends distal from ligation. The thermal cycler was obtained from Coy Laboratory Products, Ann Arbor, MI.

Following LCR procedure described in Examples 4 and 5, the mixtures were analyzed as described in Example 9 using the prototype version of the IM₂ instrument (Abbott Laboratories, Abbott Park, IL).

Referring to FIG. 5, the graph displays the results obtained from performing LCR on 10⁷ molecules of the targets. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second/

EP 0 477 972 B1

second. Background signal is approximately 15 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV 8, and those values are about 40 times background signal.

EXAMPLE 12

The following sequences were determined to be specific for a portion of the E6 region of HPV type 18.

	<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to:</u>
10	LCR7A	89	TATATTGCAA GACAGTATTG GAAC	200 - 223
	LCR7A'	90	pTTTCCAATAC TGTCTTGCAA TTTA	223 - 200
	LCR7B	91	pTTACAGAAGT ATTGAATTT GCATT	224 - 249
15	LCR7B'	92	AATGCAAATT CAAATACCTC TGTA	249 - 224

EXAMPLE 13

Plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. The plasmids were those of Example 9. All of the oligonucleotides from Example 12 which were used as probes had chemical labels covalently attached at the ends distal from ligation. The thermal cycler was as described in Example 11.

Following the LCR procedure of Examples 4 and 5, the mixtures were analyzed as described in Example 9 using the prototype version of the IMx instrument (Abbott Laboratories, Abbott Park, IL).

Referring to FIG. 6, the graph shows the results obtained from performing LCR on 10^7 molecules of the targets. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second. Background signal is approximately 15 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV 18, and those values are about 80 times background signal.

EXAMPLE 14

The following sequences were determined to be specific for a portion of the E6 region of HPV type 16

	<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to:</u>
35	LCR8A	93	GTATGGAACA ACATTAGAAC ASGA	352 - 375
	LCR8A'	94	pTGCTGTTCTA ATGTTGTTCC ATAC	375 - 352
40	LCR8B	95	pATACAACAAA CCCTTGTTGT ATTT	376 - 399
	LCR8B'	96	AAATCACACA ACGGTTTGTT GTAT	399 - 376

EXAMPLE 15

Plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. All of the oligonucleotides from Example 14 used as probes had chemical labels covalently attached at the ends distal from ligation. The thermal cycler was as described in Example 11.

Following LCR procedure of Examples 4 and 5, the mixtures were analyzed as described in Example 9 using the prototype version of the IMx instrument (Abbott Laboratories, Abbott Park, IL).

Referring to FIG. 7, the graph details the results obtained from performing LCR on 10^7 molecules of the targets. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second. Background signal is approximately 10 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV 16, and those values are about 36 times background signal.

EXAMPLE 16

The attached Appendix (example 16) discloses the sequences of the invention aligned to known sequences.

EXAMPLE 16

APPENDIX

HUMAN PAPILLOMA VIRUS

ALIGNMENT of TYPES 6, 11, 16, 18, 31, and 33; with CONSENSUS SEQUENCE

The appendix lists the sequences of HPV types 6, 11, 16, 18, 31, and 33. It also shows where the sequences of this invention line up with respect to these HPV sequences. In addition, the appendix shows where other sequences, known to the inventors as of 28 September 1990, and claimed or disclosed by or unknown to others, line up with respect to these sequences.

1. Sequences and Regions Claimed by Us,

PCR = Sequences per examples 1 through 3 (only PCR1, PCR5, PCR14 and PCR15)

LCR = Sequences per examples 4 through 14 only

2. Sequences and Regions Unknown to Others and Not Claimed by Us,

PCR = Sequences designated PCR other than those above JJ

LCR = Sequences designated LCR other than those above

3. Sequences and Regions Claimed by Others;
(Italics represents antisense sequences)

AUS = International application number (Australians) PCT/AU88/00047 (WO 88/06634)

WL = International application number (Wayne Lancaster, Wayne State University) PCT/US86/00629
(WO 86/05816)

BE = European Patent Application (Belgians) 89.033834 (X= T or U)

C = International application number (CETUS) PCT/US89/03747 (WO 90/C2821)

O = International application number (Oncor) PCT/US89/O1318 (WO 89/09940)

and

4. Sequences and Regions Disclosed by Others.

S = Sarkar, F.H. and Crissman, J.D. *Biotechniques* 9:180-184 (1990) (Italics represents antisense sequences)

6 1 gTTAATAACAATCTTgGTTTAA AAAATAGGAGGG ACCGAAA ACGGTTCAACCGAAAA
 11 1 cTTAATAACAATCTTgGTTTAA AAAATAGGAGGG ACCGAAA ACGGTTCAACCGAAAA
 5 33 1 gtaAaCTATATgCCaAGTTTA AAAA AGtAGGGCGTAACCGAAA gCGGTTCAACCGAAAA
 16 1 actACAATAAT tcatGTATA AAA ctAAGGGCGTAAACCGAAA cCGGTTCAACCGAAAc
 31 1 TAATA ATAAATAT ctTAGTATA AAA AAgTAGGGAGTgACCGAAA GtGgTGAACCGAAAA
 10 18 1 attAATActTttaaCaattgTAGTATATAAA AA AGGGAGTAACCGAAAacGgtgGgACCGAAAA
 con --taataata-ta-aa-tcttag-T-tA-AAAaaag-ACGGAgtaAACGAAA-acggtt-aACCGAAAA
 C4-GCCAASTGGCTTTT
 C5-GCCAGCCCTGGCTTTT
 15 C16-CGGTTSAACCGAAAA
 C17-CGGTCGGGACCGAAAA
 C18-CGGTTSAACCGAAAH
 C19-CGGTTCAACCGAAAH
 015-ATTAATACTTTTAAACAATTGTAGTATATAAA AA AGGGAGTAACCGAAAACGGTCGGGACCGAAAA-015
 024- ACTACAATAAT TCATGTATA AAA CTAAGGGCGTAACCGAAA TCGGTTGAACCGAAAC-024
 20 S1-CGGTCGGGACCGAAAA
 S3-ACCGAAAC
 6 58 CGGTTgTATATAAA CCAGCCCTAAAAATTAGCAaACGAGGCATTATGGAAGTgAaATGCCCTCAC
 11 58 CGGTTaTATATAAA CCAGCCCAAAAAATTAGCAGACGAGGCATTATGGAAGTAAAGATGCCCTCAC
 25 33 62 CGGTgcATATaAAAGCA aNCATTTTgcagtaAGgtctGCAcGAtATGTTTCAAGCAcTgAGGA
 16 58 CGGTtAgTATA AAAGCA gACATTTTtGcaCAaAAAgAGAACTGCaATGTTTCCacAGGA
 31 60 CGGTTgTATATAAAGCAcataGTATtTgtGCAAAccTACAgacGCcATGTTtCAaaATCCTtGcAGA
 30 18 66 CGGT GTATATATAA agatgtGagaacacacCcAcaTACtatgGCgcTtTgAygATCaaCAcy
 con CGGtt-gtatataAagcag--ca-a--at--gcaaac-aqcatt-cgatgttt-aagatc--c-ga
 GCC-C4
 GCC-C5
 35 CGG-C16
 CGG-C17 C68-CCTACAGACGCCATGTTCa-C68
 CGG-C18 C71-GCAGTAAGGTACTGCAC-C71
 CGG-C19 C16-AGATCCCAACAG
 015-CGGT GTATATAAA AGATGTGAGAAACACACCACAATCATATGGCGGCTTTGAGGATCC-015
 014-CGUTAGTATA AAAGCA GACATTTTATGCACCAAAAGAGAACTGCAATTTTCCACAGCA(-014)
 40 CGGTG-S1 S2-CCGCGGAAACTCTCAGATTGTGCG-S2
 CGGTATAGTATA AAAGC-S3

6 125 GTCTGCAACgaCeATAGACCAGTTGTGCAAGAGCTTTAACTcTcTATGcAcAGgtTGCAAAATaAT
 11 125 GTCTGCAACAcCtATAGACCAGTTGTGCAAGAGCTTTAACTcTtCTtTGcACAcCtTGCAAAATTCAGT
 5 33 129 aaAACCACGAACaTtgCAtgAtTTGTGCAAGCAtTGgAgACAACtATACACAACAtTgAAcTACAGT
 16 124 gCAGCCcGAAAgTtACcAcAgTTATGCAcACAgcTGcAAACACtATACATGAAATATATATTAGAAAT
 31 128 aaGACCTcGgAAATtGcATGAACTAAcGCTCGGcAtTGgAAATAcCtAcgATGAAcTAAGATTgAAAT
 10 18 131 gCAGCCcAcAAgctAcCTTGATCTgtGCAcGGAcTGAACAcCtCactgcAAAGAcATgAAATaACCT
 con g-gacCaagaa--tTAcAt-AgtTgtGCa-ggc-tTgAA-a-atCtatgcAt-a-aTa-aAaTaaa-T
 GTCTGCAAC-AUS1 AUS7-GCAAGACGTTTAACTCT-AUS7 C74-ACACTCTGCAAAATTCAGT
 AAGACCTC-C67
 15 010-GCGACCCTACAAAGCTACCTGATCTGTGCAcGGAACtGAACACTTCACtGCAAGACATAGAAATTAACCT-010
 024-GCGACCcAGAAAGTTACcACAGTTATGCAcAGAGCTGCAAAACAACTATACATGATATAATATTAGAAT-024
 S4-CTGGGCTCTTCAATGGGTGCAATA-S4
 6 193 GtGTGTTTTGCAAGaATGCACtGACCAcAGCAGAGATtTATcCATATGCAATATAAacACCTAAAGGTc
 11 193 GCGTGTtTTCAGgAATGCACtGACCAcGcGAGAGATATATGCAATATGcCATATAgAACCTAAAGGTt
 20 33 197 GCTGTgAATGCAAAACCTtTTCcAAcGAtCTCGAGGTATATGATTTGcATTtGgAGATTTAACAGTT
 16 192 GTGTGTACTGCAAgcAACAGTTACTgCAcAgTGAGGTATATGACTTTGcCTTTCggGATTtATgCATa
 31 196 GTGTcTACTGCAAAggcCAGTTAAcAgaAACAGAGGTATTaGAtTTTGCACTTACAGATTATAcCAATA
 25 18 199 GTGTaTATGCAAgacagtaTtggahcttACAGAGGTATTtCAaTTTGCAATTAAAGATTtAttgtg
 con GtGTyAtTGCaagaa--catTgacac-a-caGAgTtAtgaaTtTGCAttTaaAgAttTaa--GT-
 AUS2-TACGTGACTGGTGCCCGTCTC-AUS2 C73-ACACCTAAAGGTG
 GC-C74 AUS3-TGAGGTATATGCACTTTTGTCTTT-AUS3
 30 C69-GAGGTATATGCACTTTTCC-C69 01-CTAAAGGTT
 C61-GAGATWTATKCATATGC-C61 02-CTAAAGGTT
 C69-ACAGTATTGGAACCTTACAG-C69 04-GATTTCGAA
 C70-CAACAGTTACTGCGACGC-C70 06-TATATGATA
 C72-GACAGTATTGGAACCTTACAG-C70 07-TATATGATA
 S5-GTGTtTTCGAGGAATGCACTGACCA-S5 08-AATACGTAAT
 35 010-GTGATATTGCAAGACAGTACTGCAcAGGTATTCAGAGGTATTtGAATTtGCATTtAAAGATtTATTGTG-010
 011-TTATTGTG
 012-TTATTGTG
 013-AATTAACAC
 017-CTAAAGGTC
 018-CTAAAGGTC
 019-GATTTCGAG
 40 024-GTGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATATGACTTGCTTTTCCGATTTTATGCA-024
 025-TTATTGTG

```

6 261 GTGTTTCGAGGCGGCTATCCATATGCAAGCGTGGCGTGGCTAGAAATCTCAGGAAAAATAACCA
5
11 261 GTGTggCGAGACAaCtCTCCCTTCGACGgTGTGCTGTGCTTAGAAATCGAAGGAAAAATTAACCA
33 265 GTATATAGAGAGGAAATCAATTTGCaAaTGTGAaactGTGTTTgcgGTTtCtATCTAAAAATTAGTGA
16 260 GTATATAGAGAGGGAATATATGCTGCGTGAJAATGCTTTCTTAATATAGTGA
10
31 264 GTATATAggGACGacacCAcCAcAGcgAGTgTGTacaaAAATGCTTAAgaTTTATTCAAAAGTAAGTGA
18 267 GTGTATAgGAGCAgTataCCcCATgcGcGcATGccatAAATGATATgaaTTTATTCTAgaaATTAGaGA
con
gT-TatagaGacgcygagatGatGcg-aTGG--aaatGtTAgaaatTctctcTaaatAgTgA
C-44CTCTGCTGCGAGCTAGGTAHAGCGH-C44 JJ1-aattaggA
C-45CTCTGTCTATATGGCGTAGCA-C45 AUBS-GTGA
C-46CCCTGCTGTGTGGTGTGCGCT-C46 S6-GT
C-47CTCTGCTGCGHAGGCTAHACSW-C47
C-48CCTGTGTGCGHAGGCTAHACSW-C48
C-49CYCTGTGNDMGCTAHACSW-C49
C56-HGAGACRGCGWWTCCATWTG-C56
C57-HGAGACRGGSWWTCCATWTG-C57
C58-HGAGACRGWWTCCATWTG-C58
C59-HGAGACRGATATACCCGAT-C59
GTCTGGCGGAGACAACCTCTCTTGCAGCGGTGTGCTGCGTGTG-01
GTGTGGGAGACAACCTTCCC-02
03-CAACTTCTCCCTTCGACGGTGTGCTGTC-03
CACACGGCTCTGTGAAAGGGAGAACCTGCGCACAGGAGAC-04
GTATATAGAGAGGGAATCA-06
GTATATAGAGATGCGAGAGGAGCTGTGTATGCAATAAT-C07
CAGATATCTCTAGCTTAGGTATAGGACATCACTATTAC-08
09-ACCCCTTAGGTATAGCACATACATCTATTAC-09
010-GTGTATAGAGACAGTATACCCGATGCTGATGTCATAAATGATATAGATTATTCTAGAAATTAGAGA-010
GTGTATAGAGACAGTATACCG-011
GTGTATAGAGACAGTATACCCGATGCTGATGTCATAAATG-012
CACATATCTCTAGCTTAGGTATAGGAGATGATGATGATG-013
014-GTATATAGAGACAGTATACCGTATAGGTATTATAC-014
017-CTGTTTCGAGGCGGCTATCA-017
018-CTGTTTCGAGGCGGCTATCATATAGCAAGCTGCGGCGTGTG-018
019-GCCGATAGGTATATAGCTGCGGAGCGGACAGC-019
GACAAAGCTCGCGGATAGGTATATAGCTGCGGAGCGGACAGC-020
024-GTATATAGAGACAGTATAGCTGTATGTATGATAAATGTTTAAAGTTTATTCTAAAAATTAGTGA-024
GTGTATAGAGACAGTATACCG-025
026-CAGTATACCCCATGCTGATGCATGCATATAAATG-026

```

```

6 329 ATATAGACACTTTgATTATGCTGgATTGCAaCAACAGTtGAAGAAGAAACtAAcAAGAaATcTTAg
|||||
11 329 ATATAGACACTTTAAATTATGCTGcATATGCAcCTACAGtAGAAAGAAACcAAtgAAGATATtTTAA
|||||
5 33 333 ATATAGACATTATAATTATcTGTATATGGAaATACATTAGAACCAAcAgTtAAaAAACCTTTaaTG
|||||
16 328 gTATAGACATTATcGTtATAGTtTGtATGGAACACATTAGAACCAgcaAtacAAcAAACCGTtGtGTG
|||||
31 332 ATTtAGAtggtAtaGATtAGTGTGTtGGAACACATTtAGAAAAtTgACaAAcAAAGGtAtATGTG
|||||
10 18 335 ATTaAGAcattATtcAgActcTGtGTATGGAgaCAcATTgGAAAAcTtAcTtAACActGGgtTATaca
con aTatAGAcattTaTasttAt-cTgt-TATGgAacaACATtAGAA-Aa-aaactAAcaasg-t-Tat-tg
atatagacatt-JJ1
GTATAGACATTAT-AUS8
15 C50-ATAHSACAYATACSTtGtGTtMATCTT-C50
C51-ATAHSACAYATACSTtGtGTtMATc-C51
C52-ATAHSACAYATACSTtGtGTtMAT-C52
C53-CTGAGACACATACCTCTGtGTGTAACC-C53
C54-CTGAGACACATACCTCTGtGTGTAA-C54
C55-CTGAGACACATACCTCTGtGTGTAA-C55
20 010-ATTaAGACATTATTCAGACTCTGTGTATGGAACACATTGGAaaaaCTAACTAAcACTGGGTTATACA-010
024-GTATAGACATTATTGTTATAGTTTGTATGGAACAAcATTAGAACAGCAATACAAcAAACCGTtGTGTG-024
TATATCTGTGAAATTAATACGAC-S6

6 397 AcGTGcTAATTCGgTGcTACCTGTGTCAcAAaCCGtTGtGTGAAGTAGAAAA ggTAAAAcAtATAcT
25 11 397 AAGTGTtAAATTGcGTGTACCTGTGTCAcAAgCCGTtGTGTGAaATAGAAAA cTAAAGcACATAcT
|||||
33 401 AAaTaTAAATTAGGTGTATTATaTGTCAAAgaCCtTTGTGTCTcTAAAGAAAAaAaGACATgtGGAT
|||||
16 396 ATTtGTtAAATTAGGTGTATTAAcTGTCAAAagCCcTGtGTCTGTAAGAAAAgCAAGACATcTGGAC
|||||
30 31 400 ATTtGTtAAATTAGGTGTATaAcGTGTCAaagACCGTtGTGTCCAGAAgAAAAcAAAGcATtGGAT
|||||
18 403 ATTTtTTAAATaAGGTGccTgcgGTGcAGaAACCGTtGaaTCCAGcAGAAAAcTtAGACAccTtAaT
con AttTgtTAAATaGgTGtst--tgTGTCAaAaCCGtTGtGTccagaAGAAAAaca-agAcacTc--at
AUS4-IAATTAATCCACATAAT-AUS4 AUS9-GATTtATTGG
35 AUS5-TGTCATAACCTTGAMtGTCT-AUS5
010-ATTtTTAAATAAGGTCCCTGCGGTGCGAGAAACCGTtGAATCCAGcAGAAAAcTTAGACACCTTAAT-010
024-ATTtGTtAAATTAGGTGTATTAACTGTCAaagCCcATcGTGTCTGAAGAAAAgCAAGcATcTGGAC-024
40
45
50
55

```

```

6 464 aaccAAGGCGCGTTCATAAA          gCTAAATgtacGTGGAAGG          TCGcTG
      |||||
11 464 gggAAAGGCACGcTTCATAAA          CTAATAAaCcgtTGAAGG          TCGcTG
      |||||
13 469 ttAAAcAAAcGATTTCATATAT          TtcCGGTGTGtGGGCAGGCGcGTGgcGcGTGTG
      |||||
16 464 AAAAAcAAAGATTTCATAATATA          aGGGTcGGTGGACcGGTcGATGTATGtTGTG
      |||||
10 31 468 AAAAAGaACGATTCCACAACATAG          GaGGaAGGTGGACaGGaCgtTCATagCaTGTG
      |||||
18 471 gAAAAacgACGATTcACAACATAGctgggcactataGagGccagTgcattcgTcctgcaaccGagc
con  aaaaAa--acgatTtCAtAA-atag-----ctaaaagGagc-tgGcgagggcg-tgcattggct-Gttg
      TGGTGTATAGA-AUS9          AUS6-AAATGTATAGATTTTATTTC-AUS6
15 010-GAAAAACGACGATTTCACAACATAGCTGGGCACTATAGAGGCCAGTGCCATTCTGCTGCAACCGAGC-C65-CAACCGAGC
024-AAAAAGCAAGATTCCATAATATA          AGGGCTCGTGGACCGGTGCATGTATGCTTGTGTG-024

6 512          CctTACACTGC          TGGACAACATGCATG          GAAGAcTGT
      |||||
20 11 512          CttTACACTGC          TGGACAACATGCATG          GAAGAcTGT
      |||||
33 528          gaggtcccgACGTAGAGAACTGCactgtgAcgTGtAAAAacgcCATagagGACACaagcC
      |||||
16 523          cagatccacAAGAcCAGCTAGAGAAAC          CCAGcGTGAA          tCATcCATGGAGAtACAC          C
      |||||
25 31 527          GagAAGACtCCCTacTGAAGAC          CCAagTGTAA          acATcGTGGAGAAACAC          C
      |||||
18 539          acgacaGgaAcGACtcCaacgacgcAgagaascaCaAgataAtattaaGtaTgcAtggACtcaagggC
con  --ga--gagaagaccacgta-aga-Actgca---ccaggtgtAaaacatgcaTcGagagAcacaaaggc
      C64-GAACACCTTAGGAAAC          CCAG-C64
30 010-ACGACAGGAACGACTCCACGACGCAGAGAAACACAAGTATAATTAATAGTATGCATGGACCTAAGGC-010
024-CAGATCATCAAGAACACGTAGAGAAAC          CCAG-024

6 547 TACCCtTAAAGGA          TATcGTATTAGACCTGCAaCCCTcGAGCCGTGACGGTtACATTGCTATG
      |||||
35 11 547 TACCCtTAAAGGA          TATaGTATTAGACCTGCAcCCTcTGACCCGTtAGGGTtACATTGCTATG
      |||||
33 590 aACgTtTAAAGGA          ATATGtTtTAGA          TtTatATcCTGAaCCAACTGAcCTATACtGCTATG
      |||||
16 579 TAcTtTGCAtGA          ATAtaTGTtTAGA          TtTGCAACcAGaGAcCACTGAtCTCTACtGTtATG
      |||||
40 31 577 TAcgTtGCAAGAC          TAtgTGTtTAGA          TtTGCAACcTGMGcCAcTGAACCTCCACtGTtATG
      |||||
18 607 aACaTtGCAAGAccttGtAtTtgcattTAGAgccccaaaAtgnaattcGggtTGACCTtCtaGTcAcG
con  tAc-tT--AgGAc-----at-tgt-tTAGAcctt---catcc-ga-cCa--tGAcTAcacGT-tAtG
      BE16-ACCAGAGACAACXGAXCXACXGX-BE16
      BE18-GXXAGAXXGCAACcAGAGACAACXGAXCXACXGX-BE18
010-AACATTGCAAGACATTGTATTGCACTTAGAGCCCCAAATGAATTCGGGTGACCTTCTATGTGACG-010
      C89-G
      C90-G
50
55

```


6 128 CTGGTTGTGcAGTGTcACAGaAacAGACATCAGAGAAgTgCAAcAgCTTcTGTGGGcACATCAAAcAT
 11 128 CTGGTTGTGcAGTGTcACAGAcgGAGACATCAGAGAAcTACAAGAcCTTcTGTGGGCACACATAAATAT
 33 771 TtATGTGTcAaAcACTACAGcAaGtGACCtACGAACcATACAGcAaCTAcTtATGGGCACAgTgAAATAT
 16 760 TTGTGTCTACaAGcACACACTAGCATTGcAaCTTGTGAAGAcCTTtAAATGGGCACCTAGCAAT
 31 758 TTGTGTGTACAGAGcACACAGTAGAtATTGcGcAtATTGCAAGAGCTGTtAAATGGGCTCATtCGGAAT
 18 809 cTagtAGTAgAAAGCTcAGcAGcAGcGAcTTCCAGcATTCCAGcAGCTGTTtCTGaaCAcCctGtCctT
 con -Tg--tGTAcAGcGcACgAag--aGAcATtCGAcAtTgcAa--AgCTgtT--aTGGgcaCAcTaa--aT
 XGc-BE19 BE29-AGCAAGXGACCCXACGAACCAKACA-BE29 C42-CCCCTGTGAYYDDTA
 XGKXCGKXAC-BE29 C43-CTTGTGGGACAGGAA
 CXAGX-BE25
 15 BE30-AGXACAGCAAGXGACCCXAGAACCAKACAGCAACX-BE30
 010-CTAGTAGTAGAAAGCTCAGCAGAGACCTTCGAGCATTCAGCAGCTGTTCTGAAcACCCCTGTCTT-010
 6 796 aGTGTGTCCCATCTGCGC AC CgAAgaCcTAACAAcGATGGCGGACGNTTCAGGTACAGAAAT
 11 796 TGTGTGTCCCATCTGCGC AC CAaAACCAtAACAAAGATGGCGGACGATTCAGGTACAGAAAT
 33 839 TGTGTGTCCCATCTGCTGC ACaAcAAAtAAACATCAcTAcAaATGGCcGATCTCTGAAGTACAAAtGgg
 16 828 TGTGTGTCCCATCTGTTCT CAgAAAcATAATCTACcATGGCTGATCTCTGAGGTACcAATGGGGA
 31 826 GGTGTGGCCCAcTGTCTTCT actAGAcTGTAA CTAAATGGCTGATCCAGCAGGTACAGATGGGA
 25 18 877 TGTGTGtCcGtggTGTgC atCcCagCAgTAAGcAAcAATGGCTGATCCAGcAGGTACAGAcGGGA
 con tGTGTG-CCcAtCTGtGTAcAa-aacAatatacAaAAtg---Gt---g---g---tA-ag-ggat
 C40-CACACRGGGTAGACRG-C40 C75-ATGGCKGAYCTGHAGGTAC-C75
 C41-CACACAGGCACACACG-C41 C76-ATGGCKGAYGATTCAGGTAC-C76
 ACACAC-C42 C77-ATGGCKGAYCTTCAGGTAC-C77
 ACACAC-C43 CB1-TACCcGCTTGGACCTCAATG-CB1
 CB2-TACCcGCTTACCAAGTCCAATG-CB2
 CB3-TACCcGCTTACCAAGTCCAATG-CB3
 010-TGTGTGTCCGTGTGTGTGC ATCCcAGCAGTAAGCAACAAATGGCTGATC-010
 36 6 859 GAGGGGTCTGGGTGTACAGGATGGTTTATGTTAGAAAGCTATAGTgcAaCACCcAaCAGG TAC
 11 859 GAGGGGTCTGGGTGTACAGGATGGTTTATGTTAGAAAGCcATAGTAGACACACTACAGG TAC
 33 906 GctcGGAtGGGCTGTACTGGTGTGTTgAGGTAGAAGcAGTcATAGAGAgAaAACAGG aGA
 16 895 GAGGGtACGGGATGTAAATGcATGGTTTATGTAGAGGCTcGTAgTgGAAaAaAaAACAGG GGA
 40 31 891 GGGcACGGGATGcAAATGGtGTGTTTATGTAGAAGCcGTAAcTcGACAGcAgAGCAGG GGA
 18 943 GGGcACGGGtTGAaAGGGGTGTGTTTATGTAGAAAGCTcTtGTAGACAAaAaAACAGGagatgtaat
 con gagGGGgagGGGTGtA-tGGATGTTTtA-GTAGAAGCT-TagTagA-aasaaAACAGG-----a
 C79-TGTAAHNGGtTGGTTTATGT-C78
 C79-TGTAAHNGGtTGGTTTACAGGT-C79
 C80-TGTAAHNGGtTGGTTTATGTT-C80
 C84-ACATKWCCKACCAAAATACA-C84
 C85-ACATKWCCKACCAAAATACA-C85
 C86-ACATKWCCKACCAAAATACA-C86

```

6 921 ACAAAATATCAGACGATGAGGATGAGGAGGTGGAGGACAGTGGGTATGACATGGTGGACTTTATTGATG
11 921 ACAAAATATCAGAAAGATGAGGAAAGAGGAGGTGGAGGACAGTGGGTATGACATGGTGGACTTTATTGATG
33 968 TAATATTCAGAAAGATGAGGATGAAACAGCAGTACAGTGGCAGGATTtAcTAGAGTTTATAGATG
16 957 TgcTATaTCAGAtGACGAGAAcGAAAAcGacAGTGATACAGGtGAAgATTtGGTAGAtTTTATAGtAa
10 31 951 caacATTTCAGAGGACGAAATGAAAGACAGcAGTGATAcTGGGGAGGATATGGTtGAcTTTATTAGcAa
18 1009 atcagaTgacCAGGACGAAAAATG caACAG AcACaGGGtcGCATATGCTAGAtTTTATTGATa
con a-aat-tcaCA-GA-GAg-AtGaa-a-g-ggatgAcA-tGGGtagGataTggTaGACTTTATTtGat-

15 6 989 A CAGcaATATTACA CACAAATTCacTGGAAACACAGGCATTGTTTAAcAGGCAGGAGGCG
11 989 A CAGgcATATTACA CAAAAATTCcTGGAAACACAGGCATTGTTTAAAGCAGGAGGCG
33 1036 ATtcTAtgGAAAATAgTATACAGGCAGAcACAGAGGCAGCcCgGGCATTGTTTAAATAcAGGAGAGGg
20 16 1025 ATgaTAATGAtTATttaaACAGGCAGAAACAGAGACAGCACAtGcGtTGGTTTAcTGCACAGGAAGCa
31 1019 ATcGTAATGtATAcssaAAtCAAGGCAGAAgCAGAGACAGCACAGGCATTGTTTATGACAGGAGAGCg
18 1071 cacaaggaacATtcttgTgaACAGGCAGAGctAGAGACAGCACAGGCATTGTTTCATGCGCAGGAGGtc
25 con attataatgcataataataCAGgcaga--cagag-cAGCaCagGcATTGTTtaat-o-CAGGA-Gcg

6 1048 GAcACcCATTATGCGACTGTGcAGGACCTAAAAcGAAAGTATTTAGGcAGTCCATATGTTAGTCCCTAT
11 1048 GAtGCTCATTATGCGACTGTGcAGGACCTAAAAcGAAAGTATTTAGGcAGTCCATATGTAAGTCCCTAT
30 33 1104 GAgGATgATTtAaATGCTGTGtGtGcACTAAAAcGAAAGT TTGcCGc
16 1093 aAacAACATagAGATGCaGTaCAGGTTCTAAAAcGAAAGT AT TcGCTAGTCCa
31 1087 gAggAACATGcAGAGGcTcGTGcAGGTTCTAAAAcGAAAGT ATgTsGCTAGTCCc
36 18 1139 ACaAtgATGcAcAaGtGtTGCAtGTTtAAAAcGAAAGT ttgcaggagcgcagcacga
con gA-gatcATt-agaggctyTgcagGtctAAAAcGAAAGTAtttagg-agtcca--tgtga-tgCG-t
BE1-XAAAAcGAAAGX-BE1
BE2-AGGACcXAAAcGAAAGXAXXAG-BE2
BE3-AGGXcXAAAAcGAAAGXAXXGG-BE3
40 BE4-AXGXXXXAAAcGAAAGXAXXGAG-BE4

```

6 1116 AAaCacTaTAGCGgAgCCAGTgGAAAGTGAAaTAGTCCACGATTgGACGcCATTAAaCTTACAaGAC
 11 1116 AAGCAaTgTAGCTaATGCAGTaGAAAGTgAgATaAGTCCACGgTTaGACGCCATTAAaCTTACAaC
 5 11 1151 ATGctccaaAgTGCtGCGgAgGAGTGTtGAtcGTgCTgcaAacCCgtGtaqAagTcTATtAaTA
 16 1146 CTtAGTGATATTAG TGGaTGTGTaGAcAATaATATTAGTCCcTaGatTAAaAGCTATATGTa
 31 1141 tTAAGTGATATTAG TcGtGTGTGTGATtATAATATTAGTCCACGgTAAaAGCTATATGTa
 10 18 1198 aaAcagTccATTAGgggagcggtggagGTGGATacagAgTtTaAGTCCACGgTAcAAGaATATctt
 con a-aca-tatAttagagcgagtgga-gtGtggtatgt--t-taagTcc--g--taaaagctAta-gta
 15 6 1184 AGCCAAAAAAGGTAAAGCGACGGCTGTTtAAAcCaGGGAaCTAACGGACAGTGGATATGGCTATTCT
 11 1184 AGCCAAAAAAGGTAAAGCGACGGCTGTTtAAAcCaGGGAaCTAACGGACAGTGGATATGGCTATTCT
 20 33 1219 AAaAtAAAgAtGcAcAtacaGAAaAcGaaAAATAgcTGAgcTAGAAGACAGCGGATATGGCAATCT
 16 1207 TAGAAAAAaAAAGTAgAgCtGCAAAAAAGgAGAtTATTGAAagcGAAGACAGCGGGATATGGCAATCT
 31 1202 TAGAAAAATaAcAGTAAaAcCaGAAaAcGAAAGCTcTTTtAcTtcCAGACAGCGGATATGGCAATCT
 18 1266 TAAAtagTgggagAAAAagGCAAAAAAGcGcGTtTTAcAaTatCAGATAGGcGTATGGCTgTtCT
 25 con -a-aaAaaaag-g-aaa--ag-aaaa-g-a-aatatttgaacta-caGAcAG-GGATATGGC-aT-CT
 30 JJ3-tatgtctattct
 C81-ATACGGTAWGA
 C86-ATACGAYAWGA
 35 6 1252 GAAGTGGAAAGCTGgaacgggAACC CAGGTACAGAAACA TGGCG
 11 1252 GAAGTGGAAAGCTG caAAGC CAGGTAGAGAAACA TGGCG
 33 1287 GAAGTGGAAACT CAGCAGAT GGTa caACA GGTAG
 16 1275 GAAGTGGAAACT CAGCAGAT G tTACA GGTAG
 31 1270 GAAGTGGAAAC gCAGCAGAT G gTACA GGTAG
 40 18 1334 GAAGTGGAAgC aaCaCAGATtcaggtaacTACaAaattggcgaacatggcgccgaattgtatGTAG
 con GAAGTGGAA-Ctg----gca-caGatagtagag--ACA-----GtaG
 45 gaagtggaaagctghnnnnnncacagat-JJ3
 CTCACCT-C87
 CTCACCT-C88

```

6 1295          taCCGGAAATGG          GGGAGATGGTCAGGAAAGGGA
11 1289 A          cCCGGAAATGG          GGGAGATGGTCAGGAAAGGGA
33 1321 A          AagtcAAATGG          cgACAC AaActtaaatGActtaGA
16 1306 A          A          gggcGccatgagactgAAACACAtgtagtCagtAtagtGg
33 1301 A          GGAG          CAAC          AAAA          AC
10 18 1396          tgccggcagttacGGAGgctatagaCAACGggggccacagagggcAAAC          AC
con          a-----ggag-----aacgczaaatgg-----aga---gaaacacgagatggtcaggaaggga

6 1329 CACAGGaAGGCACATAGAGGG          GgaGGAAGATAcAGAGGCGGAAGGccccacaaACAGtgtAc
11 1323 CACAGGcAGGCACATAGAGGGTgagGgGGTGGAAcATAgAGAGCGCAAGCagtagacGACAGGaacCC
33 1358 atCtAGTGGGgtgGgGAtGaTccaGAAcTAAGctGTgagacaaatGtAGaTagctGTGAAA
16 1349 tggAAGTGGGgtGgttgcaGTCagtAcsgTAGTGGaagtgGGGAGagGgtTgtTAGTGAAGAGAcAc
20 33 1317 AttAAGT          tgtaATgGTAGTG          ACGGGA          cAcATAGTGAAGcGgaGga
18 1445 A          gcagtgtagacggTacaAGTG          AC          aAtAgcaaatAtaGaaAat
con          a-caagtaggacagaga-ggt-aggga-g---agtgataga-cgg--gaagca---agtgaaga-a--

6 1391 GgGAGCATGCAGgcCACAgCAGGAATAT          TgGAATTgtTAAATCTAAAGATtTAC          GggGagCATT
11 1391 GgGAGCATGCAGAGAGATcAGGAATAT          TAGAATTAcTAAATGTAAgGATATAC          GAtCtAGATT
30 33 1420          atgttACgttgCAGGAA AT          TAGTAATGTTCTAcAtAgTAGTAATACAAAAGCAaAtAT
16 1417 cTatAtgcCAaACACCaacttacAA          ATATTTTAAATGTaGTAAAAAaCTAGTAATGCAAAAGcAGCAaAT
31 1361 atGAAAcTcCCACAC          GAA ATATATTYcAAgTGTAAAAAaCTAGCAATTgtTAAAGcttGCTAT
35 18 1487 gTAAAtCcaCAAtgtaccataGcAcAttTAAaagActTGTAAAAgtaAaCAATaaacAAGGaGCTAT
con          gtgaa--t-caa-c-ca-caggaAtAtattagaaatggtt-tAaaaa--ag-saTacaagaagcgc-at

6 1455 ActTGCTAAGTTTAAAGaATGCCTTGGGCTGTCTTTTATaGATTAAATTAGGCCATTTAAAAGTGATA
40 11 1455 ACaTGGTAACTTTAAAGAcTGCCTTGGGCTGTCTTTTgtTgATTAAATTAGGCCATTTAAAAGTGATA
33 1478 ATTAtaTAAATTTAAAGAGgcCTATGGAaTAgCTTTTAtgGAATTAGTAAGACCATTTAAAAGTGATA
16 1484 GTTAGcAAAAATTTAAAGAGTTATAcGGGCTgAGTTTTTcAGNATTAGTAAGACCATTTAAAAGTAATA
45 31 1422 GTTAGgtAAATTTAAAGaATTTATATGtGTAGTTTTTAtGGAActTAAATTAGGCCATTTAAAAGCAATA
18 1555 GTTAGcagTATTAAAGAcacATATGGGcTAtcaTTTTAcAGattTAgTTAGaaatTTTAAAGTgATA
con          -ttaggtaaaTTTAAAGA-t---TatGGgcT---tTTTataGA-tTA-TAG-ccaTTTAAAGTgATA
JJ4-ttaggttagaccatttaaagtata

```

6 1523 aaACACATGTTtaGATTGGGTGGTtagCAGGgTTTGGTATACATCATAGCATATcAGAGGCATTTCAA
 11 1523 gAAcACATGtgCcGATTGGGTGGTtGCAGGATTtGGTATACATCATAGCATAGCAGAGcGATTTCaA
 33 1546 AAACAAGcGTaCcGATTGGGTATaAcCAGGATaTGGAAATtAgTCCatcagTAGCAGAAAGTTTAAAA
 16 1552 AAtCAAGcGTtgcGATTGGGTATtGCTGCATTtGGAcTTACACCcAgtaTAGCTGAcAGTtAAAA
 31 1490 AAAGCAcAGTACTGATTGGTtGTtAGCTGcGTTGGAGtTACAGgtACAgTtGCAGAAAGGATTTAA
 18 1623 AAACCAcGTATCaGATTGGgtTAcAGCTataTTGGAGtTaAacccaCAaTaGCAGAAAGGATTTAA
 con aAac-AcaTGTacaGATTGG--t-tagC-ggaTtTGGaaT-satccta-aATagCaGaggatTtAAA
 aaacaacNtgtNcagattgg-JJ4
 6 1591 AAaTTAAATtGAGCCATTAAgTTTTATATGCACATATACAATGGCTaACAAATGCATGGGGAATGGTATt
 11 1591 AAgTTAAATtGAGCCATTAAgTTTTATATGCACATATACAATGGCTtACAAATGCATGGGGAATGGTATt
 33 1614 gtATTAAATtAAACAgcATAGTTtGTATaCTaCTtTACAATGTtTAACTtGcGataGAGGAATaATaT
 20 16 1620 ACACtATtAcAAcAATATtGTTTATaTTtAcAcAtcCAAAAGTTTAGCATGTtTcATGGGGAATGGTtGt
 31 1558 ACcCTATTtCAACCATATTGTTTgtATTtCCATtAcAAAGTTTAGCATGTtTcTGGGGAATGGTtAT
 18 1691 ACACtAAATaCAgCCATTtAtaTtATATgtCCATtTcCAAtGTtTAGaCTGTtAAATGGGGAATtATaT
 con eca-TAAATtCA-Ccat-tagtTTaTATgtCAcAtaTaCAATGt-Ta-catgtgtcatGgGGAATtTgTaaT
 6 1659 gTTAGTATTATtAAAGtTTAAAGTAAATAAaAGTAGaaGTACCGTtGCACGtACAcTtGcaACGcTAT
 11 1659 ATTAGTATTAAATAAgTTTAAAGTAAATAAaAGCAGATGTACCGTtGCACGtACATtAGGTACGTTAT
 33 1682 ATTATtTgTTAAATtAGATTtAGgtGTAGCAAAACAGgttAACAGTaGCAAAAcTaaTgaGTAAATTTAT
 16 1688 GTTACTATTAGTAAAGATaTAAATGTGGAAAAATAGAgaaCAAAATTGAAAAATtTgcTtCAAACTAT
 31 1626 GTTAAATtGCTtGTGAGATtTAAATGTGcAAAAATAGAAaTAAcCAATTGAAAAATtATTAGaaAAATAT
 35 18 1759 aTTAgccCTgtTgcGtTAcAAATGTGgttAAGAgTAGAcTAACAgTTGctAAAggtTTAAgtACgtTgT
 con -TTAgatTa-TaaGaTtAaatgt-gtAAaA-tAGa-taACAgTtGcAaa-tatTaggtA-gtTaT
 6 1727 TAAATATACCTGAAAAcCAaATGTTAAATAGAGCCcCAAAAAATACAAAGTGGTgtGtgcAGCCCTCTAT
 11 1727 TAAATATACCTGAAAAcCAcATGTTAAATAGAGCCcCTCAAAATACAAAGTGGGtAGAGCCCTCTAT
 33 1750 TATCAATACCTGAAAcATGTATGtTATAGAGCCcCAAAAAATACGAgccAaCAAGtGCATTCTAT
 45 16 1756 TATGTgtTCTcCAAtGTATGtTATAGAGCCcCAAAATTCGtAGtACAGAGCAGCATTATAT
 31 1694 TgtGTATATCTCAAAATGTATATGTTAAATCAGCCACCCAAATtACGTAGcACAGCTGCAGCATTATAT
 18 1827 TaccagTAcCTgaAAcTGTATGTTATTCaCCACCAAAATtGCGAGtAgTgtTCAGACATATAT
 50 con TatatATaCTGaAAAttgtATGtTAAT-gAgCCaCCAAATTaCgaAGtagag--gcaGCa-T-TAT

6 1795 TGGTTTCGtACAGGtAtATCAAAATGccAGTACAGTTATAGCGGaaGCaCCaAATGGATAAcACGCCA
 11 1795 TGGTTTAGgACAGGcAtTtCAAAATGCaAGTACAGTTATAGGGAGgCCgCGAATGGATAAcACGCCA
 5 33 1818 TGGTTTAGAAcAGGcAtTtCAAAcATTAGTGAtGTacAAGgtacAAcACCCGAATGGATAGatAGAct
 16 1824 TGGTATAAaJACAGGtAtATCAAAcATTAGTGAAgTgTATGGaCacACgCCAGAATGGATAcAAAGACA
 31 1762 TGGTAcAGAAcAGGAATgTCAAAcATTAGcGatGTATATGGtGAAcACCCAGAATGGATagAAAGACA
 10 18 1895 TGGTAcAGAAcAGGAATtCAAAcATTAGTGAAgTAAgtGGaGACACcCTGAgTGGATAcAAAGAct
 con TGGT-tagAACAGgaATATCAAAattattAGTgaaGTaa-AGG-gaaacACCaGaaTGGATA-aaaGcCa
 BE32-AXAXCAAAAXXAGXGAAX-BE32 JJ6-tggataNaagaca
 15 6 1863 aACAGTTATTGAAcAcgGgTTGGcAGACAGTCaGTTTAAATTAACaGAAATGGTCAGTGGGcGtATG
 11 1863 gACcGTtATTGAACATAGctTTGGTCAGAGTCAATTTAAATTAACtGAAATGGTCAGTGGGcATATG
 33 1886 AAcTgTTTTACAAcATAGcTTTAAATGAAATtATATTGcAtTTAAgTGAATGGTCAGTGGGcATATG
 20 16 1892 AACAGTATTACAAcATAGTTTAAATGAtTgTACATTtGAAATTAACaCAGatGGTACAAATGGGcTAcG
 31 1830 AACAGTATTACAGCATAGCTTTTAAATGAcAcaACATTTGATTTGTCCcAAATGGTACAAATGGGcATATG
 18 1963 tAcTtAtTAACaaCATgGaaTagATGAtAGcAatTTTGTATTGTcGaaAAATGGTACAAATGGGcATcTG
 con aACAgTt-TacAAcATagcttTt-stGA-agttaaTTTgA-TTA-cagAAATGGTACA-TGGGcATatG
 25 aacNgttatacaacatgcttNgatgat-JJ6
 6 1931 ATAATGAcATtTGcAGgAGAGTGAAAtGcATTGTGAATATGCACAAaGgGAGATTTTGATtCtAAAT
 11 1931 ATAATGAtAtTtTGcGaaGAAAGTGAGAtGcATTGTGAATATGCACAGCGTGGAGAcTTTGAcTccAAAT
 30 33 1954 ATAAcGAgTtAacgGACGATAGTGAcATTGCATATtAtTATGCACAAcTTGcAGAtTCaAATAGTAAAT
 16 1960 ATAAATGAcATAgTcACAGATAGTGAATTCATATAAATATGCACAAATGGcAGAcCTTAATAGTAAAT
 31 1898 AcAAATGAgTtAtGcATGATAGTCAAAATGcCTATATAATATGCACAAATTAGcTCAGAGTgATAGTAAAT
 35 18 2011 ATAATGAgcTgAcAGATGAAGcGATAtTgGCaTtTgAAATATGccttATTAGCAGACAGCAAcAGcAAAT
 con AtAAATGA-ataa--GA-GATAGtGAaAttGcAt-TgAAATATGCacatt-GcAGc--t-AtagtAAAT
 40 6 1999 GCaGcAGcCaTTTTTAAATAGcAAATATGCAGGcCAAAATATGtGAAGAGATTGTGCACaATATGTGtAGACA
 11 1999 GCaAGGcGcTTTTTAAATAGTAAATATGCAGGcCAAAATATGtGAAGAGATTGTGCACaATATGTGcAGACA
 33 2022 GcTgCTGCATTTTTTAAAGTAAcTCACaAGCAAAATAGTAAAGGAcTGTGGAATAATGTGTAGACA
 45 16 2018 GCaAGTGGcTTTcTAAAAAGTAATTCACAGGCAAAATcGTAAAGGATTGTGcAACAAATGTGTAGACA
 31 1966 GCaCTGTCATTTTTTAAAGTAAATTCAGAGGCAAAATAGTAAAGATTGTGGAATAATGTGTAGACA
 18 2099 GCAGcTGGcTTTTTAAAGCAATTCGCAAGCAAAATattTaAAAGATTGTGccCAAAATGTGcAGaACA
 50 con GCa-gtGC-TTTTAAAAAGtAAATtcgAGGcCAAAA---gtGAaAGatTGTGCAaCAaATGTGtAgACA

6 2067 TTATAAACATGCGAATGAggAGATGYCTATaAAACAATGGATaAAacATAGGGGTCTAAAAaTAg
 11 2067 TTATAAACATGCGAATGAAaAGATGTCTATaAAACAATGGATaAagtATAGGGGTACTAAAGTtG
 33 2090 TTATAAAAAGCAGAAAAAagTcAAATGTCAATagGACAAATGGATAcAAagTAGATGTGAaAAaAcAa
 16 2096 TTATAAACGAGCAGAAAAaAaCAAAATGagTatGagTcCAATGGATAAAAATaTAGATGTGAaAagGTgag
 31 2034 TTATAAACCGCAGAAAAACGCAAAATGccatGGACAGATGGATtAAaagTAGATGTGAaAAGTta
 18 2167 TTATAggCGAGCccAAAAACGCAAAATGaatATGcACAGTGGATagcAttTAGATGTtcaAAAAaTAg
 con TTATAaac-aGCagAAAAa-ga-AaATGtctATyagaCAaTGGATaaaataTAGATGTg-tAaa-tag
 JJ11-tggataaaaatatagctgHctaaaatag
 6 2135 AagGcacAGGaAatTGGAAaCCAAATGTGTaCAaTtCtAcGACATCAAAATaTAGAAATtATTCctTTT
 11 2135 AcaGTCTaGGTAAcTGGAAgCCAAATGTgCACTTtCTAAAGACATCAAAACATAGAATTTATTCcATT
 33 2158 ATGATGGAGGAaATGGAGaCCAAATaGTaCNGTtGTTAAAGATATCAAAACATtGAATTTAcagCATTT
 16 2164 ATGATGGAGGTGAATGGAAgCAAAATtGtTatGTTTTTAAAGTATCAaAggtgTAGAGTTATGTTCATTT
 31 2102 gTGACGaAGGTGactGGAGGgaCATAGTAAGTTTTTAAAGATATCAaCAAAATAGAAstTtgTGTCATTT
 18 2235 atGAAgGgGGAaGATGGAGaCCaATAGTgcaATTCcTgGcATAcCAACAAATAGAGtTTaTaaCATTT
 26 con atgatggaGG--AtTGGa--ccaAT-GTAcagTTt-TaaGatAtcCAaa-sTGAaTTtat--CaTT
 atgatggaagaattgga-JJ11 JJ12-catbt
 6 2203 TTAActAAAtTtAAATATTAGGCTGCACGGtACGCCaAAAAAAaTcATcTAGCCATaGTAGGcCtCC
 11 2203 TTAAGcAAAcTtAAATATTAGGCTGCACGgaACGCCaAAAAAAATtGTATAGCCATTGTAGGgCCACC
 33 2226 TTAAGTGCATTtAAaAagTTTTTaaAGGTATACCAAAAAAaagGTGTtgcTAATTTgTGAACAGC
 16 2232 TTAaCTGCATTAAaAaGaTTTTTgCAAGGcATACCTAaAAAAAATtGcATaTtAcATATAGGTGCAGC
 31 2170 TTAaCTGCATTAAaAaGaTTTTTgCAAGGcATACCTAaAAAAAATtGcATaTtAcATATAGGTGCAGC
 18 2303 TTAAGgaGCCTTAAaAtcaTTTTTAAaAGhaCCcCAaAAAAAATtGTtTAgTAtTctgTGAACAGC
 con TTAa-tgcAtTaAAttttTttT--AaGGaa-gCCaAaAAAAa-TGtTaTgtaA-t-t-tGg-cCa-C
 ttaagtgcattaaaattatttttgcaggNacNcNaaaaaaa-JJ12
 6 2271 aGAcACTGGGAaTCGTACTTTTGTATGAG TTTAAATaAgcTTTcTaGGaGGACAGTTATTAGTCaT
 11 2271 tGAeACTGGGAAGTCGTCTTTTGTATGAG TTTAAATaAGTTTtTgGGgGGAACAGTTATTAGTTAT
 33 2294 aAAACAGGGAAGTCATaTTTTGCAATGAG TTTAAATaAGTTTtTaaAGGCTgTGTATATcaTgt
 16 2300 TAaACAGGTAAATCAATtTTTGTATGAG TTTAAATGAaTTTtTgCAAGGCTGTAAATAgTtT
 31 2238 TAATACAGGTAAATCAATtTTTGTGAATGAGcTTATGTAGCTTtTAcAAGGATgTAATATCATAT
 18 2371 aAATACAGGAaAATCATATTTTGGAAATGAGcTTAT acacTTTATACAGGAgcagTAATATCATAT
 con --aATAcAGG-AAATCaTatTTTgGaATGAG-tTTAataaacTTTtTacaagGatc-gT-ATat--taT

6 2338 GTAAATTCaGcAGCCATTTtGGtTgCAaCCgtTAgtaGATGCTaAggTgACATgTTgAGATGATGC
 11 2338 GTAAATTCcGcAGCCATTTtGGcTACaCCgCacTaaCgGATGCAAAAGTgGCATtATgGATGATGC
 33 2361 GTAAATTCtAAAGcTcATTTtGGTTgCAGcCATtATcAGATGCAAAATAGGAATgTAGATGATGt
 16 2367 GTAAATTCtAAAGcCATTTtGGTTACaCCATTAGCAGATGcCAAAATAGGtATGTTAGATGATGC
 21 2306 GcAAATTCaAAAGTCATTTtGGTTACaCCacTgGctGATGCTAAATAGGCATGTTAGATGATGC
 18 2438 GtgAATTCcActAGTCATTTtGGTTgAAcCgtTaaCaGATaCTAAggTgGCATGTTAGATGATGC
 con GtaAATTCcasaAG-CATtTTtGGtT-cAAcCAtTAgcaGATgCTAAa-TaG-aaTgtTgATGATGc
 6 2406 aACACaGcCATGTTGGaTATATATGGATACATATATGAGAAAtTgTTAGATGGTAATCCTATGAGtA
 11 2406 cACACaAACCATGTTGGACATATATGGATACATATATGAGAAAcctTaTTAGATGGTAATCCTATGAGcA
 33 2429 aAcGcCAAtAaGTTGGACATATATAGATGATtACATGAGAAATGcGTTAGATGGAAATgAAATTCaA
 16 2435 TACaGtGcCCTGTGGaAcTACATAGATGAGaAtTtAAGAAATGcATTgGATGGAAATtAGTTTCtA
 31 2374 TACAAACGcCATGTTGGcATATATAGAcAAAtTaccTAcGAAATGcACTAGATGGCAcCCTGATCTA
 18 2506 aAcGACcAcGtGTTGGaCaTAcTtTGATaCCtTAtaTgaGAAATGcGTTAGATGGCAATCCaATAgTAA
 con aACaccgccatGTTGGaCaTAtaTaGATA---tAtaTgaGAAAtgc-tTACATGG-AAccc-a---tA
 JJ15-gtggacatatatNgetacNtatatgagaatgcggttagatgg-JJ15
 6 2474 TtGAcAGAAAGcATaAAGCATtGACATTAATTAaATGTCACcCTcTgCTaGTAAGcTCAAcATAGAt
 11 2474 TAGATAGAAACATAGAGCATTAACATTAATTAaATGTCACcCGcTAcTgTTAcATCAAAATATAGAc
 33 2497 TAGATGTgAAACATAGGcCATTAGTgCAAtTAAaATGTCACcACTgCTtTTACcTCAAAATAcAAAT
 16 2503 TgGATGTAAAGCATAGAcCATtGtACaAcTAAaATGcCCTCCATTAATTAATTAACATCAAcATTAAT
 31 2442 TaGATGTAAAGCATAAAGCtTTAAATgCAgtTAAaATGTCCTCCcTtATTgATTACATCTAATATAAAAT
 18 2574 TtGATAgAAAGCacAAAcCATTAATaCAAcTAAaATGTCCTCCaATAcTaaCcaCAAAATATAcAT
 con TaGAT--aaAGCATa-agCATTaastacaa-TaAAATGTC-CCAcTAcTa-TtACATCAAAcATaAaAT
 6 2542 ATTAcATAAAGaAGaGATAAATaATAGtATTACATAcTAGAGTaCAcCAATTAAGATTCCAAATCCATT
 11 2542 ATTAGCAAGAGcGAGAAATAcAAATATTACATAGTAGAGTtAcACACATTTAACATTCCAAATCCATT
 33 2565 GcAGGCACAGAcCTAGATGGCCATATTACATAGTAGATTaACaGTATTgAATTtAAaAAATCCATT
 16 2571 GcLGGTACAGATTCTAGgtTGGCcttATTTACATaATAGATTGGTGGTtTTACATTTCctAAATgagTT
 31 2510 GCAGGTAAAGTgAcAGATGGCCATAcCTACATAGAcAGAcTGGTGGTtTTACATTTCcCAAAATCCATT
 18 2642 cCAGcCAAGCATaATAGATGGCCATAtTAgAAAGAGaATaacaGTATTgAATTTCcCAATgCATT
 con gc-ggtAaGAtgatAgATgGCCATAtTAcATAgTAGA-TaacagttTTAcATTTCcCAATCCaATT

55

6 2876 AGGCACATaGgaaTgCAAGTAGTgCCACCATTAAAgGTGTcCgAagCaAAAGGACATAATGCCaATTGA
 11 2876 AGGCACATcGggtTAcAAAGTAGTAcCCACCATTAAcctGTGTcAGAGACTcAAAGGACATAATGCCaATTGA
 5
 33 2902 tCACATtTatGCCACCAGtGGTGGCcttctTTGttacCATCAAAAGACcAAAGCATTTCAGTCAATTGA
 16 2908 AAACATATTAACCAACAGcAGTGGTCCACaCaeTggctGTATCAAGaataAAAGCATTACAGCAATTGA
 31 2847 CAcAgTATTAAACCACCAGGTGGTGGCAGcGtTtCaGTATCAAAAGGccAAAGCctTACAAAGCATTGA
 10
 18 2982 CAgAcATaAAACCACCAGGTGGTGGCAGCctataaacaTtTCAAAaagTAAAGCacataAAGCTATTGA
 con aaccataTaa-ccacCA-GTgGTGCCa-Cattgac-gtaTCAaAgactAAAGCa--t-AgctATTGA
 J118-tcaagactaaagcacataaagcAttga
 15
 6 2944 AATGCAAAATGCATTTAGAATCattAttAaggACTgAGTATaGTaTGAACCGTGGACATTACaAGaAA
 11 2944 AATGCAAAATGCATTTAGAATCctTTAgcAAAAACTCAGTATgGTgTGAACCTTGGACATTACAgGAcA
 33 2970 ACTaCAAAATGgCaTTAGAGaCATTAAgTAAATCACAAGTATAGTAcaaqcCAatGGACATTgCAACaAA
 20
 16 2976 ACTGCAACTAACGTAGAAACaATaATAACTCAACATATAGTAATCAaaAgTGGACATTaCAAgAcg
 31 2915 ACTaCAAAATGAtGTtggAAACaTtAAATAACTgAAATACAAAAATGAGGAcTGGACaATGCAgCaAA
 18 3050 ACTGCAAAATGgcccTaaAagggcTtgcacAAAgTcgATACAAAAccGAGGAtTGGACACTGCaAgAcA
 25
 con AcTgCAAAATg-c-tTAgAaacatTa---aaaactca-TAtagta--gaaca-TGGACAT-TCAgA-a
 actgcaaatgg-JJ18
 6 3012 CAAGTTATGAAATGTGGCaAACACCACC tAAACGcTGTtTTAAAJAAcGggGcAAaACTGTAGaAGT
 11 3012 CcAGTTATGAAATGTGGCTAACACCACC CAAACGgTGTtTTAAAJAAcAGGGAAJAtACTGTgAGGt
 30
 33 3038 CAAGCTTAgAGGTGTGGCTtTgGAACCACC AaaATGTTTTAAAJAAcAAGAGaAAcAGTaaactGT
 16 3044 tTAGCCTTGAAGTGATATTAACTGCACCAAC AGGATGTaTAAAAAAACATGGATATACAGTgAAcGT
 31 2983 CAAGcCTTGAACGTGATTtTAACGTCACTAC AGGgTGTTTTAAAJAAcCATGGATATACGTAGAGT
 35
 18 3118 CATGcgagGAACTaTggaatACaGAACCTACTcactgctTTTAAAAA ggTGGccAAcAGTAcAAcGT
 con caaG-t-tGAa-TgTggctaac-gcACCAacaa-g-tggtt-AAAAAacatGGA-A-AC-GTAgaaGT
 40
 6 3079 LAATTTTGA TGGCTGTGCaAaCAATacAATGGATATGTGGTATGGACAgATgTATgTGCAGG
 11 3079 AATTTTGA TGGCTGTGAAGaCAATgtAATGGAGTATGTGGTATGGACAcATATATAcTGCAGG
 33 3105 GCAATATGA caatGACAAaAaAATACAATGGATTATACAAACTGGggtgAaATATATATTaTG
 16 3111 GCAGTTGATGG agCATATgCAATGCAATGATGATCAAAAGTGCACaCATATATATTATTGTG
 45
 31 3050 GCAATTTGATGG tGATGTAACAACcACcATGGcATTATACTACTGGaAAtTTATATAcTATGTA
 18 3185 atAtTTTGTATGCaacaAaGacaAttgtAtgAcctATgtayCatgggaAggtgtTatTataTgacTg
 50
 con gCAaTtTGTggaacagatgaaaacaataacaAtgAttat-GaaactggacagatatATaTg--tg

6 3144 ACAATGACACeCTGGGTAAGGTgcaTAGTatgGTAGATGcTaAGGGcTaTATATTACATCTGGACAA
 5 11 3144 ACAACGACTcATGGGTAAGTAAGTACTAGTtccGTAGATGcCAAGGGcATATATTATACATCTGGACAA
 33 3170 AgCAAGAtaCATGtACTATGGTtACaGGgAAAGTAGATTTATaAGTATgTATTATATACATaAetGt
 16 3176 AAGAAAGatCagtaACTCTGGTtAGAGGgtCANGTtACATATaAGGTtATATTATCTtCATGAAGGA
 10 31 3115 TAGATtGGccaaTgTACTGTtCTGGAAGGgCAAGTTAATTgTAAGGGcCattTATTATCTACATGAAGGA
 18 3253 atGcaGGaaCATGggcaaaaccGctacctgtGTAAgTcaAGGGGatTgTATTATGTAAAGGAAGGg
 con a-gaaGacacatgg-cta-ggt-g-t-gt--aAGTagattstaagUGTaTaTATTAT-tacatgaaggga
 15 6 3212 TTTAAACATATTTATGTAAACTTTgttaaaAGGCAgAJAAAGTATGGgagCACCAaCATTTGGGAAGT
 11 3212 TTTAAACATATTTATGTAAATTTTAAAtAAAGAGGCACAJAAAGTATGGTAgTaCCAAcCATTTGGGAAGT
 33 3238 gaAJAaggtAJATTTTaaAtATTTTAAAGAGGATCCGGAAGTATtCTAAACaCaAAtgTGGGAAGT
 20 16 3244 atAcgAACATATTTTGTgAgTTTAAAGAtGATGCAGAAAAATATaGTAAAAaTAAAgTATGGGAAGT
 31 3183 cAtaAACATATTTTGTAAAtTTTACAGaAGGCAaAAAAATATGGGAtcgtTAAJAAaATGGGAAGT
 18 3321 tAcAcACgTtTTTaTaTAghaTTTAAAggtGAatgtgAJAAATATGGGAAcAcaggtACgtTGGGAAGT
 25 con t-taaaacaTaTT-TgtaaAtTTTaa--aaAGggcagaAAA-TATgg--Aa-ac-aaaaa-TGGGAAGT
 6 3280 ATGTTATGGCAGCAGAGTTAT ATGTTCTCCTGC ATCTGTATCTAGCACTaCaCAGAAGTAT
 11 3280 ATGTTATGGCAGCAGAGTTAT ATGTTCTCCTGC ATCTGTATCTAGCACTgtaCAGAGAAGTAT
 30 33 3306 ACATGtGGGTGGTCAAGTAAT tgTTGTCTCTAC gTCTATATCTAGCA ACCA AATAT
 16 3312 cATGCGGTGGTCAAGTAAT atTaTGTCCTACATCTGTgTTAGCAGCA ACGA AgTAT
 31 3251 gCATGCGGTGGTCAAGTAATTG TTTCTCTgAATCTGTaTTTAGCAG TGACGA AATAT
 36 18 3389 ACATtttGGgaaTaAtCTAATTGatgtgaaTgactctatgtgcaagTACAG TGACGACaggtTAT
 con acaT---GGT-gt-agGTaTtg-at-tt-Tcctgacdc-tct-t--c-AGcactgac-aagaagTAT
 BE211-CGGTAT
 40 6 3342 CCATTCCTGAA tCTACTACATACACCCCGCACAGACC tCcACCOCT tGTGTCTCtCaAGC AC
 11 3342 CCATTCCTGAA CTTACTACATACACCCCGCACAGACCacTgcCCCTaCaGTGTGCGctGc AC
 33 3362 CCACTaCTGAAACTgCTgACATACA gACAGAC AAC gaTaaCCGACC AC
 16 3371 CCTCTcCTGAAAtTATTAggcagCA cttgGCC AACcACcCCgCCgCGACCcATAC
 46 31 3310 CCT tTGTGgGATTCCTaAaAGCTACcaacaGCC AACcACACCcCCAcTGAATTC
 18 3454 CC qcTAcTcaGctTGTTaAaAGCTAC AgcACACCCcCCACcGtAATTC
 50 con CCact-cTgaaa---ttgacatacAccccagcagcagacc--c--caaaac-cctGC-cAacc-atac
 CC---GCXACXcAGCXG-BE21

```

6 3403 CA                      AGGAAGACGcaGTGcaaaCGCCGCCCTAGGAAaCG
5 11 3406 CAC                      GGAAGACCGcGTGctggCGCCGCTAGGAaGCG
33 3410 CACAAGC                      agcggccAAACgacGACGAC      cTgCAGacACCA
16 3426 CA AAGCCgtCGCTTGGGCACC      GAAG      AAACacacGACGAC      TATcCAGCGACCAAGA
10 31 3368 CA AAACctgCGCTTGGGCACaGtGAAGgtgTgCGGcgggcGACGACGTCTACTaAGCGACCAAGA
18 3503 CA gcACCgtgtCCgTGGGCACC      GcaAAGaccTACGGC      caGACGTc
BE10-GGCCXGXCGCGCCGCCXAG-BE10

con      CACaaaccgtgcgcttgggcacc-g-gaagcggtacgaagac-gacgacgtcc-cc-ag--accaaca

15 6 3439 AGCACGagggagtccacaCATGCCctTgCAACgCctTGTGTGTGGGCCaCATtgGacCCGTGGACAGTg
11 3442 AGCACG      tggAACGTTCaCtaacAACaCCcTGTGTGTGGGCCaCATCaGatCCGTGGACAGTg
20 33 3449      CAGACACCGCCAgcCCCT      tacaAgtGTtTGTGCA      gaCccgCCTGGACaATA
16 3481 tCAGAGCCGACACACCG      GAAACCCCTgCCACaCCaetAAGTTGTTCGAcagaGACTCaGTGGACAGTg
31 3435 sCAGAGCCGACGacC      aGAAACcCCcACcACCCaCaaAGTTGTTCGAGCGGACCTcGTGGACAGTg
25 18 3548      GgtgtgtACaCgaCCTggACaCtgtggAcTcGcgGaGaaGcagCattGTGGAC      cTg
con      acaaaagccaga--ccgc-aaaCCcct-c-acaccatgt-tttgtgtgcacagcggtctcgtGGACagTg

30 6 3507 gAAaCCaCAACcCTCATCACTAAC      AATcAGcACcAGcACCAAA      GAcGG      AACMAcAg
11 3504 cAAtCaACAACATCGTCACTgAC      AATtACAACAAGcACCAAAA      GAaGG      AACMACTG
33 3506 gAAcAgcaCgtACTGCAACTAACTGC      aCAAAcAAAGcAGCGGA      cTgtGTGT      AgTTC
16 3548 cTCCAaTcCtCACTGCattTAACAGCT      CACACAAAggACGGA      cTaaCTGT      AaTag
35 31 3502 TCAACTgtggggTtATCaGTGCAGCT      gcAtgCACAAAcCAAAcAA      GGgCTGTCAgtTcTcc
18 3604 TCAAC      ccactTcTcgGTGCAGCTacacctacaggaacAAcAAaagacGGAaaactCTcGTaGTgg

con      c--aac-c--cactgc-actaaCagctaatt-c-aacaagcacc-a-aa-----gggtgtcaaca-t-g

40 6 3562 TaaCAGTtCAGCTACGCCCTATAGTGAATtCAAGGTGaaTCCAATTGTtTAAaGTGTTTtAGATATA
11 3559 TcACAGTGCAGCTACGCCCTATAGTGAACtGCAAGGTGAcTCCAATTGTtTAAaGTGTTTtAGATATA
45 33 3561 TAAcgtTGCA      CCTATAUTGCATTtAAAGGTGaaTcCAaATAGTtTAAaGTGTTtAAAGATA
16 3603 TAAcACTACA      CCcATAGTACATTtAAaAGGTGATGCTAAATAcTTTAAaGTGTTtAAAGATA
31 3563 TgcACTACA      CCTATAATACAcTAAaAGGTGATGCAATAtaTtAAaGTGTTtAAAGATA
18 3668 TaaACTAC      gCCTATAATACAtTAAaAGGTGACagAAAcgTtTAAaGTGTTtAAAGATA

50 con      TaacacTaCagctacgCCTATAgT-CAttTaaAAGGTGAttcAaAtagttTAAaGTGTTtAAaGATa
JJ20-catttaaaaggtgAtNcNaatagtttaaaagtgttaagatat

```

6 3630 GcTgAAATGACAgacAcAgaCATTtTTGAtTTAatATCaTCAACGTGGCAcTGGGCTTcTcCaaAG
 11 3627 GaCTgAAATGACAAaTATAACATTtTGTtTGAaTTAgcATCtTCAACGTGGCAcTGGGCTTcCacCtgAG
 5
 33 3621 cAGATTtAAAAcTtTAaAgATGTGTATAGTTCaATGTCaTCCaCtTGGCAATTGGACCaGtgAcAAc
 16 3663 TAGATTtAAAAAgcATTgtacATGTATATgTgCAGTGTCTCTACATGGCAATTGGACAgGacatAAT
 31 3623 TAGgCTGtcAAAAaTAAacAATTGTATgAaCAAGTGTCACTCTACATGGCAATTGGACaTgtacagAT
 10
 18 3728 cAGATtGcgAAAAaATAgcgaccacTATaghgAtaTATCATCCaCctTGGCAATTGGACA ggtgc
 con g-agattt-aaaaa-Ata-aco-ttYTaT-a-t-a-t-TCaTC-AC-TGGCAATTGGaC-tg-cc-aa-
 gcagatt-JJ20
 15
 6 3698 GCACCACATAAA cATGCcATTGTAACTgtTAACAT ATgatAGTGAGGAACaaGcCAaCAgTTTT
 11 3695 GCACCACATAAA ATGCAATTGTAAcATtTAACAT ATAgcAGTGAGGAACAAcGtCAgCAaTTTT
 36 3688 aaaaaatagTAAAA ATGTAATTGTAACTgtTAACATtTgaAGTGAACAGCAACAAc AastgTTTT
 20
 16 3730 CtAAAAACATAAAA gTGCcAATTGTtACaCTtACATATGatAGTGAATgGCAAC GtGACcAaTTTT
 31 3690 GGAaaACATAAAA TGCTATTGTaACctTAACATATataAGTAcATCACA AGAGACgAtTTTT
 18 3791 aGgcAAtgAAAAaAcaGgaATacTgACtGTAAcATAccatAGTgAaCAcAA AGAacaaAaTTTT
 25
 con g-a-aacatAAAAaetGCAATtGTAACTgtTAACATatgatgt-aa-agcAACaaag--aacaaTTTT
 6 3762 TagAtgtTGTAAAAATACCcCCcACCATTAgcCA cAAaCTGGGAATTATGTCAcTgCAcctTATTGTA
 11 3759 TAAaCAgTGTAAAAATACCACCcACCATTAGcCAT AAGgTGGGgTTTATGTCAcTtCAATTtATTGTA
 30
 33 3752 TAGgtAccGTAAAAATACCACC tACTgTgcAAAT AAG TACTGGGATTTAT
 16 3794 TgtcTcaacCTtAAAAATACCA AAaACTaTtaCAGT GTC TACTGGATTTAT
 31 3754 TAAATACTGTtAAAAATACC tAAcACAgTatCAGT GTCaacaggatatatgactATTTA
 35
 18 3856 TAAATACTGT
 con TaaatactGTAAAAaetaccacaaacaa-tagcaat-aaggctcg-ttctatgt-actg-atttatgtga
 40
 6 3829 AtttgtatatatgttaaAtgtgtTaaATATATGTgTATgtGTGTAAacaaActgtTACaTGTATGGAAgTgG
 11 3826 A ccATTACaccTgttATATATG TATATGTGTA CATAAcATACgTGTATGGAAgTgG
 33 3801 G aCATTA TAAgTGTA CATACAAgCcaATATG
 16 3843 G tCt
 45
 31 3812 gCc
 18 3866 tgcaattccagatagtgtgcacaaatattgtgtgggataCa
 con a-----g-catta-----t-atatgtgtatatgtgtga--cataaacacatgtatggaagtcg

6 3897 TGCGCTGTACAAATGCTGCAGGAACAACCAgCAGCATTGATACT GCCTGTTaTaATTGCAT
 11 3881 TGCGCTGTACAAATGCTGCAGGAACAACCAgCAGCATTGATACT GCCTGTGTGTATTGCAT
 33 3833 TGCTGC TAAATgtATAAACCATGATAATTgTTTTTG taTTaTGTTTTaTaTT
 16 3847 atATGA caAAAcTtgatacTgcATccAcCACTTAcTgGcgTGCCTTTTG CTTTGCCTTT GGGCG
 31 3815 TAATGATTgAAcTaAatattTcTAcagtaAgCATT gTGCtaTGCCTTTTG CTTTGCCTTTTGTGTG
 18 3904 TgAcAAATgAAAcAtatgCTgTAgTaccAatATgttatCacTtaTTTTTttatTTTGCCTTTTGTGT
 con tgac-atacaa-ttgctgc-tgaacaaccca-ca--tt-ata-TgctttttggccTtt-cTtttggtt
 021-CTGCAGGAACAACCCAGCAGCATTTCATACT GCCTGTATAATTGCAT
 6 3957 TTCTtCTATGTTtTgTTAGcATcaTACTTAAtgTATggaATACCTCAgTTTaTtCTgTAcAACCTCTGCG
 11 3941 TTGcaCTATGTAATcTtTAGcATtTgTACTTAAtaTATTAAATATCTGATTTTgTaCTaTATACATCTGCG
 33 3886 gT TTTtaTgCtTaTcCtTATTATACGTCCTTTAAATAcTTTCaATTCTACcTA7gCtTggTTG
 16 3911 CT TTT GTGTGCTGcCtTATTAAATACGTCcGCTgCtTtTGTCTGCTGTACATAcaCAICATTA
 31 3880 CTactaTTT GTGTGCT tGtCATAGCTCCaCTegTgCTGTCTGTCTGgtATAtgCAaCAcTA
 18 3971
 con -tgctgtttg-tgtgt-tgcst--ta-tacgtccatt-ataatcttct-tttctgtatatacacc--tg
 021-TTGTGTATGTTTGTGTAGCATCATACTTAATGTATGGATATCTGAGTTTATTGTGTACACATCTGTG-021
 6 4025 CTaGTAATAACACTgCTTTTATATTTsCTaTTGTGGCTgCTATTAACAACCCcCTT GCAATttTTC
 11 4009 CTGGTAATAACACTgCTTTTATATTTGCTTTTGGGCTtTATTAAACAACCCcCTT GCAATtCTTT
 33 3950 CTGGT gTTGGTATTgCTgCtTtTggGtTtTGGG gAtCtCCTTTaaAAAT TTTT
 16 3974 ATAAAT ATTGGTATT acTaTTgTGGAATAAGCAgCCTCTgCgTTTaG
 31 3943 cTAtT ATTaTtgT gaTtTaTGGGtTAttGCAaCCTCTcCaTTaG
 18 3971
 con ctgtagc-tt-atcttttttataatgttcttggcttttatgaa-aac-cc-ttc-caattttt--
 021-CTAGTACTTAACACTGCTTTTATATTACTATTGTGGCTGCTATTAACAACCCcCTT GCAATTTTTC-021

6 4092 TACTAACTCTacttGTGTGtTAcTgTCCcGcTtGTATATACACTacTATATTGT taccaca
 11 4076 TACTAAcAcTgtgtGTGTGCTATATTCCTGCGCTTATATACACataTACATTGTgcaaacgcAACA
 5 33 4005 T tTGTCTATT GTTgTTTTATAT tTA CCAATGATgTgTATtAAttt
 16 4021 gTGTtTtattGTATATATatATT GTT TATA TA CCATTATTTTAAAcATAC
 31 3990 tGTGTt TGATATATAT gtTgt GTT TATA TAtattcCATTATtTgTAATtCATAC
 10 18 3971
 con tactaacactgtatat-tgctattttctgttc-ttttataactata-tccattgttttaast-cata-
 021-TACTAACTCTACTTGTGTGTACTGTCCCGCATGTATATACACTACTATATTGT TACCACA-021
 6 4154 cAgcaatGATGcTAACaTGTCAATTtAAATGATGGaGAT AcGTGcGTGgTtGTGTGTTTatgTG
 11 4144 TAAATGgTATGtTAACCTGTCTACTTtAAATGATGGaGAT ACATGGTGTtTcTGTGTGTTTtactTG
 33 4051 TCATGCACAGcatATgacacacAcaagAgTAATGTATAT ACATGtaTATATTGTtGTATATATgTG
 16 4075 ACATGCGAC gcTTTTTAattacaTAATGTATATGTACATAaTgtaATTGT ACATATA
 20 31 4044 ACATGCA tctTTTTTAA GTCAACAgTaaCTTTT AC
 18 3971 AtgcatgtatgtgtgtcTGCAtgtcccgCTTTTgcccAtctgtctgta
 con -cstgscatg-taac-t-t-Aattaaatgaagagatgacatggttg-tTtt-tg-t-t-tatgtg
 021-CAGCAATGATGCTAACATGTCAATTtAAATGATGGaGAT ACCTGGCTGGGTtTGTGTGTTATGTG-021
 26 6 4220 COTTtTaTTGATGggaTgtTgGGgTtATtTt gatGCacTAtAGcGCTGTACAGGggaTaAAC
 11 4210 CATtTgtTGTAGcTGTacttGGATTGTTGTT actaCATTacAGcGCTGTACATGGTactTgAAA
 33 4117 CA caTGGTGTgTGTtTAacATTGTTGTT gTATTTT AgtTTTTTtTtTtTGTa
 30 16 4132 atTgTGTATACCaTaacttactaTtTtcttTtTATTTTcaTataaattTTTTTTTGT
 31 4081 TtGTGTATAC TgTtTgtTgTatTgGTatttTgTatTgTgTatTgTg
 35 18 4018 tgtgtgCGTATgCAtgggtattggtatttTgtatatTgtggTaataacGTccctgcccacagcaTtc
 con catttt-tt-gtg-a-t-ttag--tt-tt-tt-tt-tt-tt-tt--a--g-t-t-tttttt-ttt-t-
 021-CCTTTATTGTAGGGAATTGGGGTTATTATT GATGCACATATAGAGCTGTACAGGGGATAAAC-021
 6 4283 AcACcAAATGTaagAAGTGTA CAAC aCAActgTAatGatGATTATGTAAcTATGcattATacT
 40 11 4273 AAACATAAGTgctAAATGTAAATCAAAcGcCAaactAcGTGTgATTATGTGTATGTGcactAGgt
 33 4171 ttACTAA TAAAT AccTTTATattttagcaGTGTAT
 16 4196 TtGTTTgtTtGTTTTTAA
 31 4124 tatttggTATtggTaTaaTaaacTtTTTACTTTTTTTAA
 45 18 4086 acagtaTATgtatTtTgttttTtTgCCaTgTtAcTattgcatatcacatgcatattgttttaca
 a-actaaactgattaaagtgttaatt-t-cc-t-t-ttT-atggtgattaagtgtatatg--tatact
 con 021-ACACCAAAATGTaagAAGTGTA CAAC ACACTGTAAATGATGATTATGTAAcTATGcATTATCT-021

55

6 4549 CAAATATTAAAAATGGGGAAGTTgGGGGTgTTTTTTGGaGGGTgGGTATaGcAcGgGtCcGcGAc
 5 11 4540 CAAATATTAAAAATGGGGAAGTTAGGGGGTgTTTTTTGGTGGGTTAGGTATTGGACACAGGgCTGGTAg
 13 4336 CAAATtCtAAATATGGcAGTTAGGGGTgTTTTTTGGTGGcTTAGGTATTGGACACAGGCTCGGTc
 16 4364 CAAATATTAcAATATGGAAGTATGGGTGTAATTTTTTTGGTGGGTTAGGAATTGGACACAGGTCgGGTAC
 10 13 4301 CAAATATTaaggTATGGLAGTATGGGTgTTTTTTGGTGGGTTgGGcTTGGGTcCgGGCTcGGTAC
 18 4358 acgtTAgcAgataAaatattgcaatGgtcaagccTTGGTataTtTtgggTGGactGGCataGGTAC
 con caaaTattaaaaatggaagttt-gGggttttttTTGGTgggTTaagttatGg-acagGcctctGGtac
 021-CAAAATATTAAAAATGGGGAAGTTTGGGGGTGTTTTTTGGAGGGTTGGGTATAGGCACGGGTTCCGGCAC-021
 15 6 4617 TGGgGTCGTaCtGGcTATgTtCCCTTcaAAcTCTgCaAAaCCTcCTATTACTaGtGGGcCtatgG
 11 4608 TGGcGGTCTGTgCaGGgTATaTACCCCTTgGGAaGcTCTCCaAGcCTGTATTACTTgGgGGGcCagCaG
 20 33 4404 AGGtGGAaGgACTGGcTATgTACCTaTtGGtActgaCCaCCtACAGCTgCAAAccCctTGCagCCtA
 26 16 4432 AGCcGACCGCACTGGgTATaTtCCcATgGGAaCAaGgCCTCCcACAGCTACAGataCAcTtgcTCTgT
 31 4369 TGGgGGTGGCACTGGaTATgTtCCcTtTaGtACAGCTCCTtCTACAGTaTCCgAGgCAaGTATACCTA
 18 4426 TGCcagTgGtCaGAGgggTgtaCagggtacattCcattgggTggggcgtTCCaAtaCAgTggTggaTg
 26 con tGGcgGtcGtaCtGGstaTgttcC-ttgggaAct-ctccc-ctacagctactaatacag-gcc-cctg
 BE11-GAAGCXKXCCCAAGCCXGXKX-BE11
 BE12-TATATACCCCTTGGGAAGCXKXCCCAAGCCXGXKX-BE12
 021-TGGGGGTGCTACTGGCTATGTTCCCTTACAAACTCTGCAAAACCTTCTATTACTAGTGGGCTATGG-021
 30 6 4685 CtCGTcctCCtGtGgTgTgTGGAGCCTGTgGGCCCTTCgGATCCaTcTatTGTGTCTtTAATTGAaGaa
 11 4676 CACGTCCgCCaGTGcTCTGGAGCCTGTtGGCCCTTCcGATCCcTcCATTGTGTCTtTAATTGAaGaaGg
 33 4472 TACGTCTcCCgGTaACtGTAGAcACTGTtGGAcCTTtaGAetCgTCTATAGTGTCTtTAATaGAAGAA
 16 4500 TAAGACCCcCcttTaAcAGTAGATCCTGTgGGcCTTtGAtCCcTCTATAGTtTCTtTACGTgAAGAA
 35 31 4437 TTAGACCAcCagTTAgCaTtGACCCTGTaGGcCTTtGGAcCCTCTATAGTAAgTcTtTGTGAAGAA
 18 4494 TTgGCCtTaCAcGTccCcaGtggTtTaTtGaaCCTgTGGgCCCCaGagaccAetCTATGTTacAtta
 con t-cGtCctcCagttac-gtaGagccTgTtGgcCctt-gGa-cCctCtatagtgctctTa-Tgaagaa
 40 BE28-CGKCCGCCGKXACGXAGAA-BE26 JJ21-TCATTTGTGTCNTTAATGAAGAA
 BE17-GKCCGCCGKXACGXAGAA-BE27 022-GGATCCATCTATTGTGTCTtTAATGAAGAA
 BE18-KCCXKCCGGKXKXAGACACGXGGACXKXAG-BE28
 021-CTCCCTCCCTCTGTGGTGGTGGAGCCTGTGGCCCTTCGGATCC-021

6 4753 TCgGCaATeATTAAcGCaGGGgGcCC TGAAaTtGTgCCCCC TgCACaCGGTGGgTTTAC
 11 4744 TCTGCTATTATTAAcGCGTGCAC TGAaggTgTAcCCCC TACACaGGGTGGgTTTAC
 5 33 4540 ACaAGTTTTATaGAGCaGGTGCACCA GccCAtTcTATTCC TACACaATCAGGtTTTga
 16 4568 ACTAGTTTTATGATGCTGGTGCACCAaCaTCTGTaCCTTcATTCCcCagatgtATCAGaGTTTTag
 31 4505 TCtGGAaTTTgTATGCTGGTGC cCCTGGCTCTTaaCCaCaCCTCTTaaCaATCTGGGTGTGA
 10 18 4562 ataGaggactccagTCGTgTtACatcaggTGCaCCTAGgCtAcgTtTAcTgGcAGTCTGGGTGTGA
 con -ctggtattatt-atGctGgtGCaCa-ctgctgc-at--c--ccccctct-caccatctGGgTTT-a
 TCTAGTNTTATTAATGCAGGTGCACC-JJ22
 BE5-CAXXAACCGAGGGGGCC-----XGAA-BE5
 15 BE6-GCAAXCAXXAACGAGGGGGCC-----XGAAAXXGXGCC-BE7
 BE7-GCAAXCAXXAACGAGGGGGCC-----XGAAAXXGXGCC-BE7
 022-TCGGCAATCATTAACGCGAGGGGGCC TGAAATGTGCCCCC TGCACAGGTGGCTTAC-022
 6 4812 aATtACATCcTCTGAaAcCaATACcCCTGCaATaTtGATGT aTcAGTT ACtAgTCACACTA
 11 4803 TATaACATCATCTGAaAcGACTACACCTGCTaTTTgATGT gTCTGTT ACaATCACACTA
 20 33 4599 TgTTACTACATCTGCAGATACTACACCTGCaATTATTAAATGTTtcaTCTGTgTgggAgTcAtTATTc
 16 4636 TATTACTACTTcaAcGATACcACACCTGC tATATTA GATattaAtAatACTGTTA
 26 31 4570 cATTgTcACaAcCtGcGAGcACaCACCTGC aTTTTTA GAT gtaACaAgTGTTA
 18 4630 tATaaCatCtGcGgtacaAcACACCTGCGgtTTTgggatctcacaccttctgtacctCtgtgtcTA
 con tATtAC--CatCtgcag--ACTaCACCTGCaAttTt-atgt--catctgtt--tac-act--ta-Ta
 05-TATACATCATCTGAATCGACTACACCTGCTaTTTgATGT gTCTGTT ACCAATCACACTA-05
 30 022-AATTACATCCTCTGAaAACCACTACCTCTGAATATTGGATGT ATCAGTT ACTAGTCACACTA-022
 6 4874 CtACTA GTaTaTTTgAAATCCcTgTtTTACAGAAcCTCTGTAAcACaACCCCAACCCcGTG
 11 4865 CcACTA GTgTgTTTcaAAATCCcTgTtTTACAGAAcCTCTGTAAATACAgCCCAACCCcGTG
 35 33 4667 aAcATTtCTACACATtAAATCCCAcTTTACGAACCTCTGTAAcACcCTCCAgGcGTGCA
 16 4692 cTACTgTtTACACACATaATCCCACTTtTACTGAGCCATCTGTATTGACAGCTCCCAACCTGCA
 31 4623 gCACACATgAaATCTTACTTTTACTGATCCATCTGTATTGACAGCTCCCAACCTGCA
 40 18 4698 tttcccaaacCAatttTaccAACTCTgCattTtCTGATCCgTCoaTtTgAggtTCCaAaACTGg
 con ctactatta-taca--TaaAAATCC-ac-TTtaCtGAaCCaCtGtTaatacAGcctCaccacCTGc-
 05-CACTA GTGTGTTTCAAAATCCCTGTtTTACAGAAcCTCTGTAAATACAGCCCAACCCcGTG-05
 022-CTACTA GTATATTAGAAATCCCTGTtTTACAGAAcCTCTGTAAATACAGCCCAACCCcGTG-022
 45 027-CTGCA

6 4939 GAGGCTAATGGACATATATTAATCTCTGCCcCcaCtgTAACgTcaCAcCctATAGAgGAAATTCCTtT
 11 4930 GAGGCCAgTGGtCacATAcTtATATCTCGCCCCaAcaATAAcATCCcAAcTgTAGAAGACATTCACt
 5 33 4735 GAAGCCtctGGaCATTtTATAcTtCTTCCTCCCTcACTgTTAGcAcACAAAgTTATGAaAACATaCCAAT
 16 4760 GAAAcTgAgGcCATtTtACAGCTTcATATCCACTATAGACAGCAATATAGAGAAATtCCAT
 17 4682 GAACcAtCAGGTcATTtAcTACTTTCATCATcAtcTATAGcACACATAATATAGAGGAATATCCTAT
 10 18 4766 GAggtggCAGGTaATgTAtTtTggtAcCccCtaCatcTgGaaACACATgggTATGAGGAAATACCTtT
 con GA-gc--c-GGtcAtTtTt-ta-TtcttTC-cC-ACcattag-acACAtaatttaGA-gAAAT-CCcAT
 05-GAGGCCAGTGGtCacATACTTATATCTCGCCCCaAcaATCCCAACATGTAGAAGACATTCACt-05
 022-GAGGCTAATGGACATATATTAATTTCTGCACCACCTGTAAcGTcACACCCCTATAGAGGAAATTCCTT-022
 027-GAAGCCTCTGGACATTtTATATTTCTTCCTCCCTACTGTTAGCACACAAAGTTATGAaAACATACCAAT-027
 15 6 5007 AGAtACTTTTGTgTATcATCTAGTGATAGcGGtCCTACATCCAGTACcCCTgTTCCTgTATGcAc
 11 4998 AGAcACTTTTGTgTATCCCTTAGTGATAGTGGaCCTACATCCAGTACcCCTgTTCCTgTgTtTtC
 20 33 4803 GGATACCTTTGTgTtTCCACAgAcagTAGTAatGTAAcATcAAGCAGcCCATTCcAGGGCTCTCGCC
 16 4828 GGATACATTtATTGTTAgCACAACcCtAAcAcAGTAACtAGTAGCACACCCATtAcCAGGGCTCTCGCC
 31 4750 GGATACATTtATTGTTTCTACTAAtaAGAAAcAaTAACAGTAGCACACCCATTCcAGGGCTCTCGCC
 18 4834 acAaACATTtTgTtTCTTgTAgggggGAggaAcCCAtAGTAGtAcCCcATgTCCcAtcTGCGCG
 25 con -gAtACATTtTgttGtttccactaatgata---aac-aAca---TAG-AC-CCcAtTCc-gg-gctcgcC
 05-AGACAGCTTTTGTgTATCCTCTAGTGATAGTGGACCTACATCCAGTACtCCTCTTCCTCTGcCCTTt-05
 022-AGATACTTTTGTGATATCATCTAGTAGAGGGCTTACATCCAGTACcCCTgTTCCTgTgTgCAC-022
 027-GGATACCTTTGTgTtTCCACGACAGTATATGTAAcATcAAGCAGcCCATTCcAGGGCTCTCGCC-027
 30 6 5075 CTCGGCCCTCGtTGggGcTAtATAGTCGTGcATTgCAcCAGGTgCAGGTtACaGACCcTGCATtTtTt
 11 5066 CTCGGCCCTCGgTGggGTTTgTATAGTCGTGcTtTaCAgCAGGTACAGGTtACGgACCcCGCgTTTtTg
 33 4871 CTCGGCCAGCCCTtGGTTATATAGTCG CAAtACcCAACAGGTtA AGTtGTgACCCCTGC
 35 16 4896 CAgtGGCAGCCCTAGGATtATATAGTCG CACaACCAACAGGTtA AAGTtGTgACCCCTGC
 31 4818 GTcctGCACGtTtTAGGgTtATATAGT AaGGCtACACAACAGTtA AAGTtATtGAtCCaAc
 18 4902 GTgTgAGcAgGTcccccGcctTtAcNGT AgGGCctacCAACAAGT gtcAGTggcTtAaCCTgA
 40 con ct-tggCAcGtct-gG-tTATATAGTgTgC-atg---a---caACAgTtAcA-gttTgttga-cctgc
 05-CTCGGCCCTCGGTGGGTTTGTATAGTCGTGCTTACAGCAGGTACAGGTtACGgACCcCGCgTTTtTg-05
 022-CTCGGCCCTCGGTGGGCTATATAGTCGTGcATTGCAcCAGGTgCAGGTtACAGACCCTGcATTtTt-022
 027-CTCTGGCAGCCCTGgTtTATATAGTCG CAAtACCCAAcAGGTtA AGTtGTgTgACCCCTGC-027
 45
 50
 55

6 5143 TCCACCTCCCAacGGGTaaTaaCATAT GAAACCCCTGTaTATGAA GGGAgAgATG
 11 5134 TCCACgCAcAagCGATTgTAACTTAT GACAACCCCTGTaTATGAA GGAAGaaGATG
 33 4932 TTTTtTAACatCgCctCaTAACTTATaACATATGATaATCTCGATtGAAAGctTtGAccttGAAG
 46 4957 TTTTtTAACatCCGcCAcTAACTTATACATATGATaAACTCCGATATGAGAGctTtAGATGtGtata
 10 4879 GTTCTCTTAgTgCTCCAAacagCTAAATACATATGAAaACCTGGCcATGAaAcTgTAAaCTtGCtCAAG
 18 4963 GTTCTCTTAcagTCCAacctcttTAAATACATATGCAaACCC gGccttTG
 con tttct-acacctaacta-taactATtAcataatGtAAaACCTGgatgatgaagt-taga-gc-gagt
 05-TCCAGCGCAcAGGAGTGTGATCTAT GACAACCCCTGTaTATGAA GSGAAGaaGTG-05
 022-TCCACTCTCAACGGCTTAAATACATAT GATAACCTGTaTATGAA GGGGAGGAGT-022
 027-TTtTtTACatCGCctCAATAAcTtTAAACATATGATaATCTCGATtTGAaAGCTtTGACctGGAAG-027
 15
 6 5198 TtaAGTgTACAaATTAgSCATGATCTA TAcCAATGCACTGATGAgGctTTATGGAaCAta
 11 5189 TaaGTTTACAaATTAcCAATAGTCTA TcCAATATGCACTGATGAAGCaTTATGATATTT
 20 33 5000 ACaCATTAACAATTtCaCATAGTAAAT TAtacccTCTCGATCTCGACTTtTcGATATTT
 16 5025 ATaCATATATAtTTTtCagTAAATGATaagTatTAtaATAGTCACAGATCCGtGATtTtGATaTa
 33 4947 AaCTTTTATCTCTTC cAtAcctCgCAATAATAGCCCTGATCCGAGACTTtTcGATATTT
 18 5013 AgcCTTgTgACaTaCattacacattTgactCtGtATgTgAtGtCtGATcAGaATTTaTgGATATT
 25
 con a-aactTtCAaAtTaaTaa-ataataatTaat-cttotaataatAGtCCGATcc-GaCTTtGtGGAaTt
 05-ATTGTTTACAaATTtACCACTAGGCTA TCCACAATGCACTGATGAAGCATTTATGATGATATT-05
 022-TTAGTGCTACAATTgATCATGATTCA TACAACATGCACTGATGAAGGCTTTTGTAGACATa-022
 027-ACaCATTAACAATTtCAACATAGTAAAT TATCACTGCTCGATCTGACTTtTGTAGATATT-027
 30
 6 5260 ATTcGtTgCAaAGACtGcCATgCGTCCgAGCctGGcCTTGCGGctAacATGcATTGGaCAAGC
 11 5251 ATTaGACTACATAGACcAGCTATaACGCTCAGaAGCGGCTTGCGGCTTTATGCGCATGGGCAAGC
 33 5062 ATTGCTATACATAGCTCTATCAaTtCTGtaGacataCTGTGCGCTTTAGTACATaGCTTCAAA
 46 5093 gTTCGCTTACATAGGCCaGCaTAACTTCAAGcCTATgGcCTATAGTaAGATGATGCTATAA
 35 33 5009 ATAGCaTACATAGGCCCTGCCctTACCtCAgtaGaaACCTGTATGATAGTAgACTAGGTATAA
 18 5081 ATcgtctATAGTAGGCTGCTcTaaCAcCAcGcGtgggACtGtTGtGctTATGATAGTATAGGTCaAag
 40
 con aTgT-tTAcATaAGCTGCTaCT-ACC-TCC-GcGtggtaCTgT-cGT-TaGTaGaAt-GGtCAa--
 JJ24-TACATAGGCTGCTATAACNCCAGNCGTGGNNITGCGNITTAGTA-JJ24
 05-ATTAGTACATACAGACAGCCTATAACGCTCCAGAGCGGCTCTTGCGCTTTATGCTCCGATGGGCAACG-05
 022-ATtGCTTtGCAcAGACTGGCATGGCTCCGACAGCTGCTGCTGGCTTTTACTAGTAGGTGCAAAa-016
 027-ATTGCTTtGCAcAGACTGGCATGGCTCCGACAGCTGCTGCTGGCTTTTACTAGTAGGTGCAAAa-022
 027-ATTGCTTtGCAcAGACTGGCTTCTATACATCTCTAGACATACCTGGCTTTTACTAGTAGGTGCAAAa-027
 028-GTACAGCTGCTGGCTTTTACTAGTAGGTGCAAAa

```

6 5328 GGGGCTCATGcACACtCGCAGcGGAAAgCacATAGGgGCCCGCAtcCATTATT
11 5319 GGGGTCcAcTtAcACACGcAGCTGGAcAACAAtAGGtCGCCCGCAtcCATTATT
5 33 5130 AGccACACtTtAAAgCTCGCAGTGGtAAACAAAttGGAcCTAGAAATACATTATTATcAgGATTtAAAGTc
16 5161 ACAAACACtAcCGtACtCGTAGTGGgAAAtcTATAgGTCGTtAgGTACATTATTATtATGATTtAAAGTA
31 5077 ACAAACtTtTGGcACtCGTAGTGGTgTcATATGGTGCAGGGTgCTtATTATTATGATATtAGTA
10 18 5149 ggcAACTTgTcttACcCGAGcGGTcAcacAAATAgGTCGTAGGGTcCAcTtTtATcATGATtAAAGTc
con -g-aaCtaTgcacACtCGcAGtGG--aaacatATAgGtGCTagg-TaCATtTtATcatgatataagta
05-GGGGTCCATGTACACACGcAGTGGACAAcATATAGTGGCCCGCAtcCATTATT(-05)
016-AGCCACACTTAAACCTCGcAGTGGTtAAACAAATTGGAGCTAGAAATcATTATcAGGATTtAAAGTc-016
15 022-GGGGTCTATGcACACtCGcAGCGGAAAGcACATAGGGGCCCGCATTcATTATT(-022)
027-AGCCCACACTTAAACCTCGcAGTGGTtAAACAAATTGGAGCTAGAAATcATTATTATcAGGATTtAAAGTc-027
028-AGCCCACACTTAAACCTCGcAGTGGTtAAACAAATTGGAGCTAGAAATcATTATTATcAGGATTtAAAGTc-028

6 5381 TttAtGAcAtTTcACcTtATTgCACAgGCTCGAGaGAAAtAGAAATGcACCCCTtTgTGG
20 11 5372 TtAgGAcATTTCACcAgTtAcACAGCTGcAGAgGAAATAGAAcTGcACCCCTcTAGTGG
33 5198 CTATTG TgcCtttAGAcCACcCGTgCcAAATgAAcAAAtTgAAAtcAGcCTtTtCaTgAtaCT
16 5229 CTATTGATCTCGAGaGAAAtagactTACAAAcTtAaAcAcCTtCTcAcATcACTACcACTTcAcAT
25 31 5145 gTATTtATCTCGAGgTGAAGTATTGAAATGCAaCCTTTAGggGcGtCTGCACTACtACTTtAcT
18 5217 cTATTgcaCCTTcCccAAATaTATTGAAcTGCAGCCTTTAG taTCTGC cAcGgag
con ctattgac-t-c---agaacacat--taca-aggct-caa--g-aac--aa-cacccc--tcg---
(05-)TTCAGGACATTTCACcAGTtACACAGCTCGAG-05
30 016-CTATTG TGCCCTTAGACcACACCGTGCcAAATGAACAATATGAATTACAGCCTTTACATGATACT-016
022- TTTATGATATTTCACCTATTGCACAGGCTCGAG-022
027-CTATTG TGCCCTTAGACcACACCGTGCcAAATGAACAATATGAATTACAGCCTTTACATGATACT-027
028-CTATTG TGCCCTTAGACcACACCGTGCcAAATGAACAATATGAATTACAGCCTTTACATGATACT-028

6 5441 CTGCACAggATGAtAcAtTTGATATTtATGCTGAaCctTTTGAaCCTgGcATtAcCCCTAcCCAACaC
36 11 5432 CTGCAGAAaATGACACGTTTGATATTtATGCTGAaCATTtTGAcCCTatCccTgACCCtGtCtAAcAT
33 5263 CtCaCaTcGtCtTaTaGTATTAAATgATGG tTtGTATGATgTTTATGC TgaCgAtGt
40 16 5297 gCAGcCTcAcCTAcTtCTATTtAAAtATGGA TtAtATGATcATTATGcAGATgacttTatAcAGA
31 5213 ttAAATGAtggCTTtTaTGAcATTtATGCAGA CAcTGATtTtAcTGGGATAcACCTgCcAcAcA
18 5273 gacAAATGA CTtGtTtGATATATATGCAGATgCAAtgGAcCTgCAgTGCcTgACCAcTgCGgttc
con -t-aa---a---tat-T-ttAt-taTg-acag-ac-atgacatt-tgc---taccctt-ccaa---
45 016-TCTACATCGCTTATAGTATTtAAATgATGG TTTGTATGATGTTtATGc TGACCACT-016
027-TCTACATCGCTTATAGTATTtAAATgATGG TTTGTATGATGTTtATGc TGACCACT-027
028-TCTACATCGCTTATAGTATTtAAATgATGG TTTGTATGATGTTtATGc TGACCACT-028

50
55

```

[illegible]

EP 0 477 972 B1

6 5713 GGTCTCTGgaTTtTATTGTCATCCTgCATGGTAATTGGACGtAaACGGCGTAAACGTATTCCCTATT
11 5695 GGTCTCTGACTTcTATTGTCATCCTACATGGTAATTGGACGcACACGGCGTAAACGTATTCCCTATT
5 33 5518 GGTCTCTGACTTgTtTTACATCCTAGTATTtTtATTTACGtcGcaGgCGTAAACGTATTCCATATT
16 5561 GcaGGTGACTTTtTATTACATCCTAGTATTAGATGTTACGaaacGACGTAACGTTTACCATATT
31 5477 GggGGTGATTtTTATTGCAcCCTAGTATTATATGTTAaaacgtCGACGTAACGTTTAcCATATT
10 18 5537 GgtacacATTATATTGtggCattaTATTATtTtaTtccataagaaACGTAAACGTGTCcCATATT
con GgtgctgacFtttaTTTgcatCCtag-TatTat-Tttttacgta-acgaCGTAAACGT-TtcC-TatTT
JJ25-CGTAAACGTTTCCCTATT
PCR2-CGTTTCCATATT
027-GGTGCTGACtTTGTTTACATCCTAGTATTtTtTtTACGTcGCAGGCGTAAACGTTTCCATATT-027
15 6 5781 TTTTcCAGATGt GGGGGcTAGcACAGCAGCATATATGTGcCTCCCTcTAACcCTGTATCC
11 5763 TTTTACAGATGt GGGGGcTAGcACAGCAGCATATATGTGcCTCCCTcTAACcCTGTATCC
20 33 5586 TTTTACAGATGTCcgtTgTGGcGGCCTAGTGAGGCCACAGTgTAcTGGcCTCCT GTAcCTGTATCT
16 5629 TTTTcCAGATGTCtTtTGGcGGCCTAGTGAGGCCACGTGtACTTGGcCTCCT GTCCcAGTATCT
31 5545 TTTTAcAGATGTCtCTGtGGcGGCCTAGcGAGGCTACTGTCtACTTACCACCT GTCCcAGTGTCT
18 5605 TTTTgCAGATGgCtTtGTGGcGGCCTAGtGacaaTAcCGTATAtcTCCACCT ccttCtGTGcGc
25 con TTTTAcAGATGtctctgtGGCgGCCCTAG-GA--cAcAGaTAT--TgCCtCTCtcc-gtccCtGTAtCt
TTTTcCAGATGTCtTtTGGcGGCCTAGTGA-JJ25
PCR1-CAGATGTCCTGTGGcGGCCTAGTG-PCR1
TTTTGcAGATG-PCR2
027-TTTTACAGATGTCcGTGTGGcGGCCTAGTGAGGCCACAGTGTACCTGGcCTCCT GTAcCTGTATCT-027
30 6 5843 AaGTTGTGTCcACGGATGCTtATGTTActCGCACAAACATATTTTATCATGCCAGAGTCTAGACT
11 5825 AaGTTGTGTCcACGGATGCGTATGTtAaCGCACAAACATATTTTATCATGCCAGAGTCTAGACT
35 33 5651 AaGTTGTcAGCACTGATGAATATGTgtcCGCACAAgCATtTATTATtATGCTGGtAGTTCAGACT
16 5694 AaGTTGTtAAGCACGGATGAATATGTtGcAGGCACAAACATATATTATCATGCCAGGAaCatTCAGACT
31 5610 AaGTTGTtAAGCACGGATGAATATGTtAaCGCACAAACATATATTATCATGCCAGGAcGtGCTAGACT
40 18 5670 AaGTTGTGAAtAcCGATGAtTATGTgAcTcCccAcAaGcATATtTTATCATGCTGGcAGcCTAGAGT
con AaGTTGTtAaGcACgATGAtTATGTtAc-CgcAC-AaCATAT-TTATcAtGC-gGAgTtCtAGACT
JJ26-GTTGTNANCACGGATGANTATGTtACTCGCACAA-JJ26
PCR3-AAGTTGTtAAGCACCGATGAATATGT-PCR3
LCR1A-AAGTTGTtAAGCACCGATGAATATGT-LCR1A
LCR1B-TGCACGCACAAACATATATTATCA-LCRB
LCR1B'-ACGTGGCTGTTTGTATATATAGTA-LCRB'
027-AAAGTGTCTCAGCACTGATGAATATGTGTCTCGCACAGCATTTATTATTATGCTGTAGTTCAGACT-027
6 5911 tCTTTCAGtGGGACATCctTatttTtCCAtaAAA cggGcTAA C AAAA CtGTGTGc

11 5893 CCTTCTGTGGGACATCCATATTactCTATCAAAAA aGCTAA C AAAA CAGTGTAC
 5 33 5719 tCTTCTGTGTGGCATTCCATATTTTCTATTAAAAAACCCTAGTAA CgCTAAAAAAATATTGTAC
 16 5762 aCTTCAAGTGGGACATCCcTATTtTCTATTAAAAAACCTAACAAAT AAcAAAA TATTaGTTc
 31 5678 gCTTACAGTAgGGCATTCCATATTtCcATTacCTAAAcCTGACAAATccTAAAAA TAGTGTgTc
 10 18 5738 attTAACtGTTGGcAATCCATATT TtaggggttcCTGcaggTgTgCgAAATAgcagGaTaTc
 con -cTtgc-GTcGAcATCCaTATTtttctaTtasaasacctgctaatt--cascasaaAaa-tagtTgTc
 JJ27-GTGGACATCCATATTT-JJ27
 027-TCCTGCTGTGGGACATCCATATTTTCTATTAAAAATCCTACTAA CGCTAAAAAAATATTGTGTAC-027
 15 6 5967 CAAAGGTGTcAGGATATCAATAcaggGtAtTTTAAGGTgGTGTtACCAGATCCTAACAAaITTGcATTg
 11 5949 CAAAGGTGTcLGGATATCAATATAGaGTgTTTAAGGTaGTGTtGCCAGATCCTAACAAgITTGcATTa
 20 33 5784 cAAAGTATCAGGcTgCAATATAGGgtTtTAGGgtCcgtTtACCAGATCCTAAATAaITTGcATTt
 16 5824 CtAAAGTATCAGGATTACAAATAcAGGgtATTtAGGaaTaCaTTTACCTGAcCCcAAATAgITTGcATTt
 31 5743 CaAAGGTgTcAGGATTACAAATATAGGgtATTtTAGGgtTcgtTTTACCAGATCCAAAcAAATITGGaTTt
 25 18 5800 CtAAGGTtTcGcATAcCAATATAGaGTATTtAGGgtTcAgTTTACCTGAcCCAAATAAATITGGcTTa
 con CAaAgGTgTcAGgAt--CAATATAGgGTATTtAggtT-c--TTaCCAGATCCTAA-AAaITTGaATTt
 JJ28-CAATATAGGgtATTtAGGgtTNCNGTTACC-28 30-AAATAAATITGGATTN
 PCR4-GTATATCCCAATAATCCCATGTAA-PCR4PCR5-TTATTTAAACCAAA
 027-CCAAAGTATCAGGCTTGCATATAGGgtTTTtTAGGgtCCGTtTACCAGATCCTAAATAAATITGGATTt-027
 30 6 6035 CCTGAcTCgTcTcttTtGcATcCCcACAAcACaCGTTTAGTATGGGCaTGcACAGGcTaGAGGTgGG
 11 6017 CCTGAtTcaTCCcTgTTTtGACCCcACTACACAgCGTTTAGTATGGGcGTCACAGGgtTgGAGGTAGG
 35 33 5852 CCTGACACCTCCCTTTATAAGCCcGATACACAGCaTTAGTATGGGCaTGcTAGGcCTGAAaTAGG
 16 5892 CCTGACACCTCaTTTATAATCCaGATACACAgCGcTgGTTTGGGcCTGTGTAGGTgTTGAGGTAGG
 31 5811 CCTGTATAcTCTTTTATAATCCGTAAAcTCAACGcTTAGTTTGGGcCTGTGTGTGTTtAGAGGTAGG
 40 18 5868 CCTGATAcagTaTTTATAATCCGTAAACaCAAGCtTtAGTgTGGGcCTGTGTGTGagTgGaaTTGG
 con CCTGA-sC-tcttTtTataATCCtGa-AcCaCaCGtTtAGTaTGGGcCTGg-sGG--T-GAgTgTaGG
 CCTGACACCTCNMTTTATAAT-JJ30
 GGAGTGTGA-PCR5
 027-CCTGACACCTCCCTTTATAACCCGTATACACAGCATTAGTATGGGcATGTAGGcCTTGAATAGG-027
 45 50 55

6 6103 CAGGGGcAaGcCaTTAGGtGTgGGTcAaAGTGGcATCCctTtCtAAAtAAATGATGATGtGtGAAA
 11 6085 CAGGGGcCaaCcTtTAGGCGTtGGTGTtAGTGGcCATCCaTtGCTAAACAAATtGATGATGtGtGAAA
 5 13 5920 TAGAGGcCAGCCcTtTAGGCGTtAGGtGtCATCCtTtATTAAACAAATtGATGATGtGtGAAA
 16 5960 TCGcGGcCAGCCcTtTAGGtGTgGGCATTAGTGGcCATCCtTtATTAAATAAATtGATGATGtGtGAAA
 31 5879 TCGcGGcCAGCCcTtTAGGtGTgGGcATTAGTGGcCATCCaTtATTAAATAAATtGATGATGtGtGAAA
 10 18 5936 cCGtGGtCAGCCcTtTAGGtGTtGGcCtTAGTGGcCATCCaTtTatAAATAAATtGATGATGtGtGAAA
 con --G-GGtCAGCCcTtTAGGtGTtGG--aTtAGTGG-CATCC-TTatttAAAtAAATtGATGATGtGtGAAA
 O27-TAGAGGcCAGCCcTtTAGGCGTtGGCATAAGTGGtCATCCtTtATTAAACAAATtGATGATGtGtGAAA-027
 15 6 6171 AT tcaGGGgagTGGTGGTAaCCCTGGaCAGGATAAcAGGGTTAAATAGGTATGGATTATAAACAA
 11 6153 ATaGTggtGGGTATGGTGGTAaCCCTGGTcAGGATAATAGGGTTAAATAGGTATGGATTATAAACAA
 20 33 5988 cCGGTAAcCaGTATcCTGGaCAaCCGGTgctGATTAATAGGGAATGCTtATCCaTGGATTATAAACAA
 16 6028 ATcGTAGTgctTATGCCaGcAaATgCaGGTGTgGATAATAGaGAATGTATATcATGGATTATAAACAA
 31 5947 AcTCTAAATagaTATGCCGgTggtTcCTGGcactGATAATAGGGAATGTATATcAATGGATTATAAACAA
 26 6 6004 gTtCCcATgcccGcCaCgtcTaaTgtTtctgagGAGctTAGGGaCAaTgtTgtTgTgATATTAAGcAg
 con at-ctaaTgggtatgctggttaatcctggtgagGATAaATAGgGaa--TgtatcttaTgGATTATAAACAA
 O27-CCGGTAACAAAGTATCCCTGGaCAaCCGGTGTCTGATAATAGGGAATGTTTATCCATGGATTATAAACAA-027
 30 6 6236 ACaCAaTATGcATGGTtGGaTGTGtGcCCcCCtTtTgGcGcAGCAITGGGGTAAaGGTAaACAgTGTaC
 11 6221 AcccAGcTATGtATGGTgGGcTGTGtGcCCcCCtTtTAGGcTGAACATTTGGGGTAAaGGTAAcACAAAGTGTc
 33 6056 ACACAGTTATGTTTAcTTGGAATGAAGCCtCCaACAGGGGGAaCAATTTGGGGTAAaGGTgTgTGTGTAC
 35 16 6096 ACACAAATGTCTTtTaatTGGTtGCAAAACCACCTATAGGGGGAACAcTGGGGAaAGGATccCAtGTAC
 31 6015 ACACAAAcTGTCTTtTaatTGGTtGCAAAACCACCTATAGGAAGAcTGGGGTAAaAGTAgTCTTGTAG
 18 6072 ACACAgTtATGtTtTtTgGcTGTgcccCtTgCTATtGGGGAaCAcTGGGcTAAAGGAcTgCTTGTAA
 40 con ACaCA-tTATGt-Ta-TGG-TGt---CCAcCtataGGGGAaCAcTGGGGTAAaAGTAcTcctTGTaC
 O27-ACACAGTtATGTTACTTGGATGTAAGCCTCCAACAGGGGAACATTTGGGGTAAaAGTGTGCTTGTAC-027
 45 50 55

6 6304 EAATACacCTGTACAggcTGGTGACTGCCGCCCTTeGAACCTTATACCAAGTGTATACAGGATGGGc
 11 6289 aAATACcctCTGTACaAaaTGGTGACTGCCGCCCTTgGAACCTTATACCAAGTGTATACAGGATGGGg
 33 6124 TAA TGCA GcAcCTgCcaTTCGATCTCCACCTTAGAACTATATAJAACTACTATTATTAGGGATGGGc
 16 6164 cAATgtTGCAGTAAATCCAGGTGATGTCTCCACCATTAGAGTTAAATAAACACAGTATTATACAGGATGGTG
 31 6083 TAACAaTGCTaTtACcCCGGTGATGTCTCCATAGAAATTAATAAAATACAGTATATACAGATGGGg
 18 6140 atcgcgTcCTTtATcCaGgGgcGATTGcCCcCLTTAGAAcTtAAAAAACaCAGTtTtTgAAGATGGGg
 con -aata-tgCtgtag--cctgggtGAtTG-CC-CC-PTaGAacTtAtaAacac-gTtATacAGgATGGGc
 027- TAATGCAGACCTGCCAATGATTGTCCACCTTTAGAACTTATAAAATACTATTATTCAGGATGGTG-027
 6 6372 AcATGGTtGAcACAGGCTTTGGTGCTATGAATTTGcCgAtTTgCagACCAATAAATCaGATGTCCct
 11 6357 AcATGGTtGAcACAGGCTTTGGTGCTATGAATTTGCGAGAcTTcAAACCAATAAATCgATGTTCcC
 33 6189 ATATGGTgGAcACAGGATTGGTtgcATGGATTtTAaAACATTgCAGGCTTAATAAAAGTAgTGTTCct
 16 6232 ATATGGTtCATAcLGGCTTTGGTGCTATGAcTTTACTACATTACAGGCTAAcAAAAGTGAAGTTCCa
 31 6351 ATATGGTtGAcACAGGCTTTGGAGGATGGAATTTACTgCtTTACAAAGAcCTAAAGTAAATGTTCct
 18 6208 ATATGGTgATAcLGGATATGGTGGcATGGAcTTTAgTAcATtTgCAAGAcACTAAATGTgAgGTACcA
 con AtATGGTtGAcACAGGcTtTGGTgctATGcGAtTTTact-cATT-CA-gc-AtaAAA-gTgATGTCCct
 ATATGGTtGATACAGGCTTTGGTGCTATGA-36 37-CATTTCIANGCAATAAAAGTGAATGTTCCT
 027-ATATGGTgGAcACAGGATTGGCTGCAATGGAATTTAAACATTCGAGGCTAATAAAAGTAgTGTTCCT-027
 6 6440 aTTGAaTatTCTGcAcTtAcTgTcAAATATCCaGATtATtTcCAAAATGGCTGCAGACCCaTATGGTGA
 11 6425 cTTGATATTtTGTGGAcCTgTcTGCAAAATATCCcGATTATTgCAAAATGGCTGCAGACCTTATGGTGA
 33 6257 aTTGATATTtTGTGcAgTAcATGCAAAATATCCAGATTATTtAATAATGcATgTgAgCCTTATGGTGA
 16 6300 cTGGATATTtTGTAcTcATTtTGGCAAAATATCCAGATTATtTtAATAATGGTgCAGAcCAATAGGGGA
 31 6219 TTGGcAATTtTGAATCTATTtTGAATATCCAGATTATcTTAATAATGGTGTGcGAGcCATATGGGGA
 18 6276 TTGGcAATTtTGTcAgTCTATTtTGTAAATATCCcGATTATtTAcAAATGcTGCAGATCCctATGGGGA
 con -T-GATATtTGTg--cTtAtTG-AAATATCCaGATTATtTc-AAAATGGcTgCA-CC-TATGcTGA
 NTNGATATTTCG-JJ37 JJ38-AAATATCCAGATTATTtAATAATGG-JJ38
 027-ATTGATATTtTGTGcGATACATGCAAAATATCCAGATTATTtAATAATGcTATGAGcCTTATGGTGA-027

6 5508 TAGaTATTTTCTtCtaCgAAGGCAAAATGTTGCGAGCAcTtTTTtTAAAGGGCGTGGgaag
 11 5493 TAGgTgTTTTTtATtTgCGAAAGGAACAAATGTTGCLGAGCACTtTTTtAATAGGGCGGTACG
 5 13 6325 TAGCTtTgTTTTTtCTtCTCGAGCGAACAAATGTTGTAGGACACtTTTtAATAGGGCTGTGTACt
 16 6328 cAGGtTATtTTTTTtATtTgTAGAGCAAAATGTTGTGACGACCTtTtTtAATAGGCGTGTACAT
 31 6287 TAcAtTAtTTTTtTtATtTgACGAGGGAACAAATGTTGTGAAGGCATTTTTtAATAGAcCAGGcACg
 10 18 6344 TtCcAtTgTTTTTtTgCTAGCGgcCTGAGCAGctTTTGctAGGCATTTTTggaATAGAcgAGGLAcTa
 con tag-tTATTTTTtAtTtTtCGaaggGAAcAAATgTTGTG-tAGAcATTTtTtAAAGGcTGGcTgcT
 WO 86/05816-GAGG
 027-TAGtTtATTTTTCTtCTCTCGACGtGAACAAATGTTGTAGGACACtTTTtAATAGGGCGGTGTACAT-027
 15 6 5576 TGGGGGAACCTGTGCCTGTAcacCTtAtTaaTtAAGGtAGTggaAAATGcAcCgTCTGTAGgYAGTAGT
 11 5561 TGGGGGAACCTGTGCCTGTAGACCTGTGgTaAaAAGGgGTtAaAaAcAGatcATCTGTAGctAGTAGT
 20 33 6393 TAGAGAGAGCGCTCTcAGTACCTGCTGCTGCTGCTcGAACTAGGCTGCTCTtTtAGAcAGT
 16 6436 TGGTGAAaATGTAcCGAcGAcTtTATACATtTAAAGGCTcTGCGTCTACTGCaAATTgAcCGAGtCa
 31 6355 TGGGTGAAtCGGtCTcTactGACTtATATATtAAAGGCTcCGGTtCaAcAGcTACTtTAGCTAaCgT
 25 18 6412 TgGGTGAacGTgCTcTactcCTtATATATtTAAAGGCaCAGGTatgCGCTcCAcctTgcAgcCTGt
 con TggGGGAa-ctGTGCTgtagaa-Ata-aTtAAAGGctGTgtgattcgtc-tct-tage-Ag-agt
 TGGGGGAACCTGTGGCTGACTACATtATATtAAAGGCTGTGGAATcCGACGTCTAGTA-W086/05816
 LCR1A-CAGCTGTGTAAGGCGTATPA-LCR3A
 LCR2A'-GGACAACCAATTTCCCCCATTAT'-LCR2A'
 LCR2B-CAGATCATCTGTAGTAGTAGT
 LCR2B'-GTCTAGTAGACATCGATCATCA
 LCR3A-ATTTATACATtTAAAGGCTcTGGTc-LCR3A
 LCR3A'-AAATGTATATTTTCCGAGACCGG-LCR3A'
 LCR3B-TACTGCAAAATTAGCCAGTTCA
 LCR3B'-ATGACGTTTAAATGGCTCAAGT
 LCR4A-CTTTATATATtTAAAGGCACAGGTAT-LCR4A
 LCR4A'-GAATATATATTTTCGCTGCTAT-LCR4A'
 LCR4B-GCTGCTGCTACCTGCGACGCTGT
 LCR4B'-CGGAGAGATGGACCGACGA
 35 027-TAGGAGAGCGGTTCGGAGTACTCTGTACATtTAAAGGCTcCAGGAcACTGCGCTCTATTCAAGACAGT-027

6 6644 ATATATGTTaAcACcCgAGcGGCTCtTTGGTGTCtCTcAGGCaCAATTgTTTAAATAgCCATATTG
 11 6629 ATTATATGtACaTAcCAGCTAGTGGCTCATTTGGTGCTTCaGAGGCTCAATTATTTAAATAAaCCATATTG
 33 6461 gcTTtTTTCCcAcCTAGTGGtCAATGGTACTTCcGaaCTCAGTTATTTAAATAgCCATATTG
 16 6504 AaTATTTTCCTACACCTAGTGGtTcATGGTTAcCTcGATGccAAATATTcAATAACCTtATTTG
 31 6423 AcATAcTTTCCTACACCTAGcGGCTCcATGGTTAcCTcGATGCaAAATTTTAAATAAcCAATTG
 10 18 6480 gtGTATtTCCcTcCTcCAAGtGGCTcATLTGTACCTcGactCcAGtTgTTTAAATAAcCAATTG
 con attTattttccc-aCaCCtAGtGGCTcAtgGTtAc-TCtGA-gC-CAattATtTAAATAAcCAATTG
 AT-LCR2B JJ39-GTTACNTCTGANGCNCAATTATTTAAATAAcCAATTG
 TAA-LCR2B'
 AA-LCR3B
 TTA-LCR3B'
 15 GT-LCR4B
 CAC-LCR4B'
 027-GCTTTTTTCCCACTCCTAGTGGATCAATGGTACTTCCGAATCTCAGTTATTTAAATAGCCATATTG-027
 6 6712 GCTaCAAAAAcCcCAGGGaCATAAaATGGTATTGtTGGGGtAAACaAcTGTTGTACTGTGGTAG
 20 11 6697 GCTcCAAAAAGcCTcAGGGaCATAAaATGGTATTGtTGGGGtAAcCAcTGTGTACTGTGGTAG
 33 6529 GCTACAAcGtGCaCAAGGtCATAAaATGGTATTGTTGGGGtAAcCAgTATTGTACTGTGGTAG
 16 6572 GtTACAAcGcGCACAGGGcCACAATATGGcATTTGTTGGGGtAAcCAATATTGTACTGTGGTAG
 25 31 6491 CaTgCAACGtGCTcAGGGaCAAAaATGGTATTGTTGGGGtAAcCAgTATTGTACTGTGGTAG
 18 6548 GtTAcAtaaGcCaCAGGGtCAcAAcAATGGTgTTGcTGGcAtAATCAaTTATTGTACTGTGGTAG
 con GcTAcAa---GcCaGgGdCaAaA-AATGGtATTGtTGGGGtAAcCAaTtATTGTACTGTGtGtG
 GCTACAAANNNGCACA-JJ39 J41-AATGGTATTGTTGGGGtAAcCAaTtATTGTACTGTGGTAG
 C6-GCHCAGGGWCAATAAATGG-C6 C1-CGTGTGGTAG
 30 C7-CTGTGTTG
 C8-CTGTGGTAG
 C10-CAGTTGTAG
 C11-CTGTGGTAG
 C12-CTGTGTTG
 C13-CTGTGGTAG
 35 C14-CTGTGGTAG
 C15-CTGTAGTGG
 027-GCTACAAcGtGCaCAAGGtCATAAaATGGTATTGTTGGGGtAAcCAgTATTGTACTGTGGTAG-027

```

6 6780 ATACCACACGGAGTACCAATGACATATG      TGCATCCTGaaCTaCTaCTTccACATACACc
5      |||||
11 6785 ATACCACACGGAGTACCAATATGACACATG      TGCATCTGTgCTTAAATCTgCTACATACACt
      |||||
13 6597 ATACCACCTGGAGTACTAATATGACCTTATG      caCacaaAGTaaACTAgTgACAgTACATATAAA
      |||||
16 6640 ATACCACACGGAGTACCAATATGCTATATGTGCT      GcCATALCCACTTcaGAAACATACATATAAA
      |||||
10 11 6559 ATACCACACGGAGTACCAATATGCTgTGTGCT      GCATTgCaacagTgATACTACATATAAA
      |||||
16 6616 ATACCACCTCCAGTACCAATeTaaCaATaTGTGCTTctacCagagTcttctgtacctgggcaATATgAt
con ATACCACaCgcAGTACCAATaTgaCatTaTGTgct---tgCa---g---aacta---ag---taactaCTatataa
      |||||
      ATACC--3741      C16-CATCCGGTAACACATCTGCA-C16
      ATACCACGGAGTAC-C1      C17-PCGTGCTAAATCTGCTACA-C17
15 ATACTACACGGAGTAC-C7      C20-CACACAAGTAACTAGTGACAG-C20
      ATACCACCTGGAGTAC-C8      C23-CAGTCTCCTGTACCTGGG-C23
      ATACTACTCGCAGCAC-C10      C31-TTGCAAAACAGTGATACTACATT-C31
      ATACTACCCGTAGTAC-C11
      ATACTACCAAGAGCAC-C12
      ATACTACTAGAACAC-C13
20 ATACCACACGGAGTAC-C14
      ACACTACCCGGAGTAC-C15
      027-ATACCACCTGGAGTACTAATATGACTTTATG      CACACAAGTAACTAGTGACAGTACATATAAA-027

6 6842 AATTCTcATTATAaGagTACATGCGcCATGTGGAAgAGTATGATTACAAaTTTATTtTTCaATTaTG
25      |||||
11 6827 AATTcAGATTATAaGAAATACATGCGcCATGTGGAgGAGTtTGATTACAGTTtATTtTtCAATTgTG
      |||||
13 6559 AATgaAAATTtTAAaGAATaTaTaaGACATGTTgAaGAATATGATcTACAGTTgTtTtTtCAACTaTG
      |||||
16 6705 AATAcTAAcTTTAAgGAGTAcTaaGACATGGgGAGGAATATGATTACAGTTtATTtTtCAACTgTG
      |||||
30 11 6624 AgTATAAcTTTAAaGAGTATtTAAGACATGGTGAGGAATcTGATTtTACAAaTTTATaTTTcAGTtTG
      |||||
16 6684 gCTaccAAaTTTAAgcAGTATagcAGACATGtTGAGCAATaTGATTtTgCagTTTATtTtTcAGTtTG
      |||||
con aaTactaATtTtTAA-gAgTA-ata-GcCATGt-GAgGAATaTGATcTaaCagTTTtTtTtCAaATt-TG
      027-AATGAAAAATTTTAAAGATATATAAGACATGTTGAAGAAATATGATCTACAGTTTGTTTtTtCAACTATG-027

35 6 6910 TAGCATTAATCTTgTCTGctGAAGTaATGGCCTATATtCACACAATGAATCCCTCTGTTtTGGAAgACT
      |||||
11 6895 TAGCATTAATCTTgTCTGAGAGTcTATGGCCATATaCACACAATGAATCCCTCTGTTtTGGAGgACT
      |||||
40 13 6721 CAAAgTTACCTTAACTGCAGAGTtTATGACATATATtCATgCTATGAATCAGaTATTtTtCaAGaGATt
      |||||
16 6773 CAAAAATAACTTAACTGCAGAGTtTATGACATAcATaCATtCTATGAATcCaaCTATTtTGGAGgAGaCT
      |||||
13 6692 CAAATTAACATTAaCTGCAGAGCaTaTACATATATGAGTATGAGTAAATcCTgATATTtTGGAGaGATt
      |||||
45 18 6752 TaCTATaCTcTTAACTGCAGATgTtATGtCcTATATTCTaAGTATGAATAgcagTATTtTtCaGAGATT
      |||||
con --AaaaTtACaTtTa-CTGCaGAagTtATGcC-TatAttCa-actATGAATccc-cTATTtTgCA-GA-T
      027-CAAAGTTACCTTAACTGCAG-027

50
55

```

6 6978 GGAACCTTGGGTTATCGCCTCCcCAAATGGTACATaGaaGATACcTATAGgTATGTgCAGTCACAG
 11 6963 GGAACCTTGGGTTATCGCCTCCcCAAATGGTACAcTgAGGATACtTATAGaTATGTaCAGTCACAG
 33 6795 GGcAaTTTGGTTTaaCACCTCCtCCAtctGtTAgttTaaAGGATACcTATAGGTTTGTtaccTcCAG
 16 6841 GGAATTTTGGTcTaaACCTCCcCCAggAGGcACacTAGAaGATACtTATAGGTTTGT aaCCcAG
 31 6760 GGAATTTTGGATgaaCaaCACCTCCcCAGGTcCTTTGGAGGATACcTATAGGTTTGTaccTCaCAG
 18 6820 GGAACCTTGGGtTtcCCcCCcCCcCCaaCtactagTTTGGtGGATACaTATcGtTTGTaCaatCTcgtt
 con GGAcTTTGGtTfa-c-cCtCCcCCaaCTggtac-tT-gagGATACcTATAGgTTGTaCa-tCacag
 C21-TTGT AACCCAG
 C34-TTGT AACCCAG
 C35-GTTTGT AACCCAG
 6 7046 GCCATTACcTGTCaaAAgCCCACTcCTGAAAAgGAAAAGCcAga TCCCTATAAGaAccTtAGTT
 11 7031 GCCATTACcTGTCagAAACCCACACCTGAAAAaGAAAAACAgGA TCCCTATAAGgAaTgAGTT
 33 6863 GCtATTACgTGTCAAAAaaCagTACCTCCAAAgGAAAAGGAAGA CCCCCTAggtAAATATaCaT
 16 6906 GCaATTgCtTGTCAAAAaCaTaCACCTCCAgcaCCtAaaGAAGATgatCCCCCTTA AAAAATAcAcT
 31 6828 GCcATTACaTGTCAAaAAacTGCcCCcCAaaagCCcAAgGAAGAT CCaTTTA AAgATtAtgtaT
 18 6888 GCtATTACcTGTCAAAaggaTGCTgCaCgggtgaaAataAgGAT CCcTaT gAtaAgTtaaagT
 con GCcATTAcTGTCaaAAacct-cacCtc-aaaggaaAggaaGAt---cCccttaaa-aA-t-ta-tT
 GCAATTGCT-C21 C18-CATACACCTCCAGCACCTAA-C18 J746-T
 C19-GGATGCTGCACCGGCTGA-C19
 C22-AAAAACAGTACCTCCAAAGGA-C22
 C27-TTTTGTcATGAGGTTTCCT-C27
 C24-CACACCTGAAAAaGAAAAACAg-C24
 C28-GTGTGACCTTTTCTTTTGTGTC-C28
 C25-CTCCTGAAAAAGGAAAAAGCCA-C25
 C26-GAGGACTTTTCTTTTGGGT-C26
 C29-CCAAAGCCCAAGGAAGAT C-C29
 C30-CAAAAGCCCAAGGAAGAT C-C30
 C32-CAGAAACCCACACCTGAAAAAGA-C32
 C33-AGAAACCCACACCTGAAAAAGAA-C33
 GCAATTGCT-C34 O23-GGA TCCCTATAAGGATATAGATT
 GCAATTGCT-C35 O15-GGAT CCCCTAT GATAAGTTAAAGT

6 7110 TTTGGGAGGTAAATTAAAGGAAAGTTTCTAGTGAATTGATCAGTATCCCTtTgGGACGcAAgTTT
 11 7095 TTTGGGAGGTAACTTAAAGGAAAGTTTCAAGTGAATTAGATCAGTTTCCTcTGGACGcAAgTTT
 33 6927 TTTGGGAGGTgagTTAAAGGAAAGTTTCTGACGAGATTTAGATCAGTTTCCTtTgGGACGcAAgTTT
 16 6973 TTTGGGAGGTaAATTAAAGGAAAGTTTCTGACGAGcTAGATCAGTTTCCTTTAGGAGcGAAATTT
 31 6892 TTTGGGAGGTtAATTAAAGGAAAGTTTCTGACAGATTAGATCAGTTTCcAcTgGGtCGCAAAATTT
 18 6952 TTTGGAAtGtTgagTTTAAAGGAAAGTTTCTcTAGATTTAGATCAATATCCCTcTGGcCGCAAAATTT
 con TTTGGGAGGTtAAtTTAAAGGAAAGTTTCTGcAGATtTaGATCAGtTTTCCT-TgGGAGcAA-TTT
 TTTGGGAGGTtAATTAAAGGAAAGTTTCTGACAGATTAGATCA-JJ46
 C2-GATCAGTTTCCTTTGGAGC-C2
 C3-GATCAGTTTCCTTTGGAGC-C3
 C7-CTAGCTCAAGGTRAMCTGC-C7
 015-TTTGGAATGTGGAITTAAGGAAAGTTTCTTAGACTTAGATCAATATCCCTTGGAGCTAAATTT-015
 023-TTTGGGAGGTAACTTAAAGGAAAGTTTCAAGTGAATTAGATCAGTTTCCTTGGAGCTAAATTT-023
 20 6 7178 TT gTT aCAAAGTGGATATgGGAGCGGTcCT
 11 7163 TTA TT gCAAAGTGGATATcGGAGCGGAcGt
 33 6995 TTA TTAcagGcAggtctttaaagcAaaAcctaaacttaaACctgcAGcccccaCAtCcaCCcgCA
 16 7041 TTA cTACAGCAGGATgaaGcGcAaACCAAAATTAcAttaGgaAAAGcAaaagCTaCACCcGcA
 25 31 6960 TTA tTACAGGcCAGGATatAGGGcCagctCCTAAATTTAAAGCAGGtAAAGC TagTgCACCC t
 18 7020 TT ggtTcAGGcTGGATtgcGtcgcaagCccAcCaTaggcCtCgcAAAGC T tctg
 con TTA--ta--agcaggatttgagggcAaaacCAaaataaa-a-cagcAaaa-gatatag-gcacc-cct
 015-TT GGTTCAGGCTGgATTTCGCTGCAGGCCACCATAGGCCCTCGCAAAAGC T TCTG-015
 023-TTA TT GCAAAGTGGATATCGGAGCGGAGCT-023
 35 6 7209 CTatTCGTACAGGTgTAAAGCGCCCTGCTGTtTcAAAGcCTCTgCtGCCCCtAAAGCTAAAGCGGcgCC
 11 7194 CTgCTCGTACAGGTaTAAAGCGCCcAGCTGTgTctAAAGcCTCTTACAGGCCcAAAGcGAAAGcGTACc
 33 7059 CaTgTCTgCAaaacgcAAaaaggttaaaaaAATAAcAcTtGtgtAattgtgtATgtTGTgtTtTg
 16 7108 CcTCATCTACcCTcTCAACAcGcTAAAGC cAAAAAGcTAAAGcGTAA GtATGTATGTA
 40 31 7021 CagCATCTACCACTACAcCaGcAAAGcGtAAAAAAC TAAaaAGTAAtgatgTCTATGTAAataCaT
 18 7075 CtcCATCTgCCATAC gctctc TAAA ccTGCcAagCgt
 con Ct-catcTaC-actscAaa---c---a--t-sat-aa-gtaa-ctg-a-cC-ct-a-c-tgta--tcc-
 015-CTGCATCTGCcACAC CcTCTc TAAA CcTGCcAGCGT-015
 023-CTGCTCGTACAGGTATAAAGCGCCcAGCTGTGCTTAAGCCCTCTACAGCCCCAAAGcGAAAGcGTACC-023

```

5      6 7277 AAAACCAAAGGATGAATATATGCTGT      aTaGTacGTGT
          |||||
11 7262 AAAACCAAAAAGATGAATATATGCTGTcgcGTGTGTGTGT
          |||||
33 7127 TctGTctATcGTactTtgtTGT      TGCTTGTGTgtTGT
          |||||
16 7167 TgtTGaaTaGTGT      TGTGTGT      TGCTGT      ATATGT
          |||||
31 7088 GTGTctgtatGTcTATGTGCTTGTgtgtatGT      ATATGT      GTGTGTGTgttgtgtgtTATATA tg
          |||||
10 18 7113 GTG      cGTGTAcGTGC      caGgaagtAATATGTGTGTGT      gtatatATATAcata
          |||||
con      -t-tctataaagtgat-tggttgtg----tgtGtagtgt-tatgttgttgt-----tatata---
015-GTG      CGGTGACGTGC      CAGAAGATATATGTGTGTGT      GTATATATATACAT-015
16 023-AAAACCAAAGATGAATATATGCTGTGTcagGTGTGTGTGT-(023)

6 7313      AT
          |||||
20 11 7302      AttTatATG      T      TGTGTGA      gGTGT
          |||||
33 7167      TTGT      TcttctGTATG      TgttaccatGtatGtTATGTTGTACGTacGTGTGTG
          |||||
16 7199      TTGTATGTgtctGTATG      TG      CTGTGAAATATTAAgtTGTATGT      GTGTGTG
          |||||
31 7150 gTATATGTATGTTTATGTATG      CgtGTGT      acTTGTATATAT      GtatagTATGT      TATGTGTGT
          |||||
25 18 7164 cTATctgtgtGCTT      GTATGcCtGTGTgttgtgtTGT      TgtAT      gTtgcattGTATG      G
          |||||
con      -kattgtatggttttgtatg-c-ttgtgt-tgt-cttgtatatattatgtgttgttgt-gtgggtgtg
015-CTATGTGTGTGTGT      GTATGCTCTGTGTGTGTGTGTGT      TOTAT      G      ATTGCATGTGTATG      G-015
30 023-      ATTATATATG      T      TGTGTGA      GTGTGT-(023)

6 7315      ATATATGT
          |||||
11 7325      aTATGT      TcTTGT      atTGTG      TATATGT
          |||||
33 7221 T      ttTATGTgtTactGTgttGTGTGcATGTtTcTATGTacttgt
          |||||
35 16 7248 TATGTATG      gtaTAATAAA      caeGTGTGTATGT
          |||||
31 7209 TATGTATGctatgatGTGTAATAAAatgtgtataccgtgtgtgtGTGTGTATGTGTcctTataTAc
          |||||
18 7221 TATGTATG      gtGTGT      gtGTATGTGTATGTATGTacTat
          |||||
40 con      tatgtatg-----tgtaataaa-----ttatgt-ttcttgt-tgtgttatgt-tatgta-tac
015-TATGTATG      GTGTGT      ATATGT      TcTGTGT      ATGTGT      TATATGT-(023)
023-

```

6 7323 GT gTATGTAcTGT
 11 7351 GTataTGTtTGTGTATATGT GTAT gtTATGTa TGT
 5 33 7262 cagttccctgttTGTGTATATGTtataaaacattgtTgtGTATtTgtTaaActATtTgtTATGTa TGT
 16 7279 GTtTtaAaTgcTtTgttAACTATTGT gTcATgcAAcATaAaTaaactaatT
 31 7277 AccctaTtagtaacatacTatTAcTAtTtataAACTATTGTtccTActTgtTcctAcTtTgttCCTTgc
 10 18 7257 AtttgtTgggtatgtggcaTtaaAraAaaTatgttttTgtgtTctgtTgtTtaTgtggTtTgcgcCCTag
 con a---tat-tgtttgtgtatat-ataataaagaacatgttcttttatgtataatt--Tatgtactgt
 015-ATTGTGTGGTATGTGGCATTAAATAAAATATGTTTGTGGTTCGTGTRGTATGTGGTTCGCCCTAG-015
 023- GTATATGTTTGTGTATATGT GTAT GTTATGTa TGT-023
 15 6 7336 TATGT aTATGT GTGTGTGTGTtctGTGTGaaTgtAgtTAITTTGTGtAAATGTGTATGTGTGT
 11 7386 TATGTgtTATGTatGtTGTGTGTtTaaGTGTGT gtATaTATTGTGGAAATGTGTATGTATGT
 20 33 7329 TATGT AtatgggtgtaccTataTGaGTaaGagTtGTATTGcTtGccctacCcTGCATTGc
 16 7331 gTTCaacAcctACtaaTgtTgtTgtTgtTtTcaTGTATaTaaactaTatTtGctACATGcTgtTt
 31 7345 TccTCCcaAtagtCATgtTatTtTctgccTataAATTAggTgTcacgccataGTaAAAGTgtgtaca
 25 18 7325 TgagtaacAactgtATtTgtTgtTgtTgtgggtTtTgcttTgtTgggctatataTgtccTgtatatt
 con tatgttaa-aa-gt-atttgtg-ctttT-tgtgtgttaattttatttgc-t-aa-ctgtatgt-cttt
 015-TGAGTAAACAAGTATTTGTGTTCGTGGTATGGCTGTGTCTGTGTGGCTATATATTCTCTGTATTT-015
 023-TATGTTGTATGTATGTGTGTGTGTATTAGTGT GTATATATTGTGGAATGTGTATGTATT-023
 30 6 7400 TaTGTGCAATAAACAATTAcctcTtgtTcacCCCTGT gACTCAGTGgctgttgccagcGTTtTGgt
 11 7450 TcTGTGCAATAAACAATTA TtatgtgtgtcCTGTACACCCAGTG actaAGTgtTgtT
 33 7390 aaTGtaCcTaccTttATtccctTaTatTgtAGtaCCTACATGTttaGtattgCtttacCTTTGaca
 35 16 7399 ttgtTTtaTataTactaTatTtTgTAgcgcCAGgcCCatTTTGTaGcTtCaAcCgaAttCggTTGcat
 31 7413 CccGgTccgtTtTtgcAACTaaAgctacTCCATTtTgaTTTtatGCagcCcatTTTAAATcccaAAC
 18 7393 CaaGtTataaaacTgcacACcttAcagcaTCCATTTtatccTacaatccCcaTtTgtcTgtgcAACCC
 40 con tatgttcгаа-aatt-atdacCttata-t-lcc-tt-t-acat-cagtg-c-atttta--cggtt-act
 015-CAAGTATATAAAACGTGACACCTTACAGCATCCATTTATCCTACAATCCTCCATTTCGTGTGCAACC-015
 023-TTTGTGCAATAAACAATTA TTATGTGTGTCTGTACACCCAGTG ACTAAAGTGTGT-023
 024-GAATTCGTTGCAAT
 45 50 55

6 7466 TTGCACGGCCCTTaccacscataagTaATATacaTgcAcaATATATATATtTtttqtTaaasTACTAT
 11 7508 TTGCACGGCCCGTtTgtttgtccCTGCATAT TaaatTATATATATtTgTaaatTAcCTATACATATg
 5 33 7458 TactAgTgTcCaaTATtTgtacaaTTTCcTccattTgTATGcCTAaccgTtTtcggTAACTTgCCATac
 16 7467 GcTtTtTgTgCaaAaaTgTgtTtTtTaaAagTtTCTATtTcCagaaCTATgTtTtTaaCTTgTACGT
 31 7481 GTTTTCGGTTGCATtTgTtTaaacaTgctAgTACaaCTATGctgtacgtaGRTTGCggTtTtTgGZ
 10 18 7461 GcTtTTCGGTTGC ctttggcTtTaTcctgTggTtTt
 con -ttt-cgg-ccttat-t-ta-a-ttc-tataa-t-ctatgt-tatat-ttt-tt-t-actttgct-tt
 015-GATTTCGGTTGC CTTTGGCTTATGTCGCGGGTtTt-015
 023-TTGCACGGCCGCTTTTGTGTTCCTTCA7AT TATATTATATATATTGTAAATATACCTATACATATC-023
 024-GCTTTTGGCACAAAATGTGTTTTTAAATAGTCTATGTGACGCAACTATGCTTTAAACTTCGTACGT-024
 15 6 7533 aCttttatAtTTGCAACGGTTTTCGGTTGCCCTTAgCATACACTTcCaaCeAAITGTTAcAAc
 11 7573 tTACCCcccceCacTTGCAACCGTTTTTCGGTTGCCCTTA CATACACTTAcCTCaaAAITGTTATAAc
 20 33 7526 aTACCCtaTgaCatTGGCagaacAgTtaaTccTtTtCTttCTGCACTGtTgTgTcTgTACTTgctg
 16 7535 TTCCCTG cTtTgCcaTgCgtTgCcaAaTcccTgtTtTCTgCCTGACCTG cTtTgccaAcaTtcc
 31 7549 TTCCCTG aaTAcTgTTTTtGCCaacaATTCTggcgTgTAgT
 25 18 7496 cTgCacaatcacgtacgctgtgcactattgcaacctTAACTtTTTggCactgCTcTacaTatTtTg
 con tt-c-ct-tt-catt-gcagcc--tttg-tt-ctcttctc-T-cact-c-tcttct-tattata-c
 015-CTGCACAAATACAGTACGCTGCGCACTATTGCAAACTTTAACTCTTTGGGCACTGCTCCTACATATTGTTG-015
 023-TTACCCCGCCGCACTTGCACCGTTTTTCGGTTGCCCTTA CATACACTTACTCaaAAITGTTATAAc-023
 024-TTCCCTG CTGCGCACTGCGCAAAATCCCTGTTTTCTGACCTGCACCTG CTGCGCAAGCATGCC-024
 30 6 7597 GTGTTTccTctTAACTCctATATattTTGTG CeAGGTACaATTCGCCCTGCCAAGTgCTTGCCAA
 11 7640 GTGTTTtTgTACTAATCCATAT gTGTGTgCaaAGGTACaATTCGCCCTGCCAAGTgCTTGCCAA
 33 7594 caTTggcaTACaAcCctATgacattTgCagaaCagTtAACTctTtCTTTcCTgcacTgtTtTgtc
 35 16 7598 aTtTgttTtTACActgCacTatgtgcaACTactgAACTAcTatTgTcATTtTgTcTataTAAaaTaaAT
 31 7589 tTCTgctcaTACACacCTTgccaATATATccAgTCCaaCTtTGCaaATTATaCTATgAACTatgT
 18 7564 aaCaattggcgCgCctCTTtggcgCATATAA ggCgacCTGgtATTA gtcATtTctCtgcg
 40 con -t-ttt-tt-ta-ca-tccttat--tt-taa-cdca-g-aca--Ttgc-tt-c--aatt-ttt--A-
 015-AACAATTGCGCGCCCTTTCGGCGCATATAA GCGGCACTGCTGTA GTCATTTTCGTGTC-015
 023-GTGTtTTCGACTAATCCCATAT G-023
 024-ATGCTTTTTTACACTGCACATGTGCAACTACTGAATCACTATGTACATGTGCTATATAAAATAAAT-024

[illegible]

25
Claims for the following Contracting States : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, NL, SE

1. A composition useful in LCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:

LCR5: SEQ ID No.

81	GCTGCAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA;

LCR6: SEQ ID No.

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCAT	GTCTTGCAGT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG;

LCR7: SEQ ID No.

89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGCTTGCAA	TITA,
91	TTACAGAAGT	ATTGTAATT	GCATT,
92	AATGCAAAAT	CAAAATACCTC	TGTAA, and

LCR8: SEQ ID No.

93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTCTA	ATGTTGTTC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT,
96	AAATCACACA	ACGGTTTGT	GTAT

2. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 16 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:
LCR5 (SEQ ID Nos. 81, 82, 83 and 84) and LCR8 (SEQ ID Nos. 93, 94, 95 and 96).
3. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 18 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:
LCR 6 (SEQ ID Nos. 85, 86, 87 and 88) and LCR 7 (SEQ ID Nos. 89, 90, 91 and 92).
4. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising:
a composition according to any of claims 1 to 3; and
further comprising a ligase.
5. A kit according to claim 4, wherein said ligase is thermostable.
6. A composition useful in PCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising

a first nucleic acid primer of sense direction, capable of hybridizing to the antisense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

1 CAGATGTC TC TGTGGCGGCC TAGTG,

6 GAATTAGTTA GACCATTTAA AAG,

7 GGGGAAACAC CAGAATGGAT A,

81 GCTGCAACA ACTATACATG ATATAA,

85 CTTCAC TGCA AGACATAGAA ATAA,

89 TATATTGCAA GACAGTATTG GAAC and

93 GTATGGAACA ACATTAGAAC AGCA; and

a second nucleic acid primer of antisense direction, capable of hybridizing to the sense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

5 AGGTGTCAGG AAAACCAAA TTAAT,

84 TGCTTGCA GT ACACACATTC TAATA,

88 TACTGTCTTG CAATATACAC AGG,

92 AATGCAAA TT CAAATACCTC TGTA and

96 AAATCACACA ACGTTTGTG GTAT;

provided said first and second primers hybridize to their respective antisense and sense strands at locations such that their 3' ends do not overlap and, in the direction of extension, the 5' ends of said primers are spaced further apart than the 3' ends of said primers.

7. A composition according to claim 6 wherein said first and second primers are selected from the following pairs of oligonucleotide sequences (identified by Sequence ID No.):
1 and 5, 6 and 5, 7 and 5, 81 and 84,
85 and 88, 89 and 92, and 93 and 96.
8. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising:
a composition according to claim 6 or 7; and
further comprising a polymerase.

9. A kit according to claim 8 wherein said polymerase is thermostable.

10. A consensus oligonucleotide for hybridizing human papilloma virus types 6, 11, 16, 18, 31, 33 and 61, which oligonucleotide comprises from about 10 to about 60 nucleotides in length and is selected from the group of sequences consisting of:

SEQ ID No.

1	CAGATGTCCTC	TGTGGC8GCC	TAGTG,
5	AGGTGTCAAG	AAAACCAAAT	TTATT,
6	GAATTAGTTA	GACCATTTAA	AAG and
7	GGGGAACAC	CAGAATGGAT	A;

and their complements.

11. A type-specific oligonucleotide for determining the presence of human papilloma virus type 16, having a sequence selected from the group consisting of:

SEQ ID No.

81	GCTGCAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTT8TT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA,
93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTCTA	ATGTTGTTC	ATAC,
95	ATACAACAAA	CCGTGTGTG	ATTT and
96	AAATCACACA	ACGTTT8TT	GTAT;

and their complements.

12. A type-specific oligonucleotide for determining the presence of human papilloma virus type 18, having a sequence selected from the group consisting of: SEQ ID No.

SEQ ID No.

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GTCTTGCACT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG,
89	TATAATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGCTTGCAA	TTTA,
91	TTACAGAGGT	ATTGAAATT	GCATT and
92	AATGCAAAAT	CAATACCTC	TGTAA;

and their complements.

13. A method for determining the presence of any human papilloma virus in a test sample, comprising

a. hybridizing DNA in the test sample with at least one consensus oligonucleotide selected from the group of claim 10, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b. determining the presence of human papilloma virus by detecting the signal generated

14. A method for determining the presence of human papilloma virus type 16 in a test sample, comprising

a. hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of claim 11, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b. determining the presence of human papilloma virus by detecting the signal generated.

15. A method for determining the presence of human papilloma virus type 18 in a test sample, comprising:
- hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of claim 12, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and
 - determining the presence of human papilloma virus by detecting the signal generated
16. A method according to any of claims 13-15, further comprising a step of amplification prior to or concurrent with said hybridizing step.
17. A method according to claim 16, wherein said amplification step comprises PCR or LCR.

Claims for the following Contracting States: ES

1. A composition useful in LCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets

LCR5: SEQ ID No.

81	GCTGCAAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCA3C,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA;

LCR6: SEQ ID No.

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GCTTGCAGT	GAA,
87	CCTGTGTATA	TGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG;

LCR7: SEQ ID No.

89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGTCTTGCAA	TTTA,
91	TTACAGAGGT	ATTGGAATTT	GCAATT,
92	AATGCAAATT	CAAAATACCTC	TGTAA; and

LCR8: SEQ ID No.

93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTCC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT,
96	AAATCACACA	ACGGTTTGT	GTAT.

2. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 16 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:
- LCR5 (SEQ ID Nos. 81,82,83 and 84) and LCR6 (SEQ ID Nos. 93, 94, 95 and 96).
3. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 18 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:
- LCR6 (SEQ ID Nos. 85,86,87 and 88) and LCR 7 (SEQ ID Nos. 89,90,91 and 92)
4. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising

a composition according to any of claims 1 to 3; and further comprising a ligase.

5. A kit according to claim 4, wherein said ligase is thermostable.

6. A composition useful in PCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising:

a first nucleic acid primer of sense direction, capable of hybridizing to the antisense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

1 CAGATGTCTC TGTGGCGGCC TAGTG.

6 GAATTAGTTA GACCATTTAA AAG.

7 GGGGAAACAC CAGAATGGAT A.

81 GCTGCAACA ACTATACATG ATATAA.

85 CTTCACTGCA AGACATAGAA ATAA.

89 TATATTGCAA GACAGTATTG GAAC and

93 GTATGGAACA ACATTAGAAC AGCA; and

a second nucleic acid primer of antisense direction, capable of hybridizing to the sense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

5 AGGTGTCAGG AAAACCAAAAT TTATT.

84 TGCTTGCACT ACACACATTC TAATA.

88 TACTGTCTTG CAATATACAC AGG.

92 AATGCAAAAT CAAATACCTC GTAA and

96 AAATCACACA ACGTTTGTG GTAT;

provided said first and second primers hybridize to their respective antisense and sense strands at locations such that their 3' ends do not overlap and, in the direction of extension, the 5' ends of said primers are spaced further apart than the 3' ends of said primers.

7. A composition according to claim 6 wherein said first and second primers are selected from the following pairs of oligonucleotide sequences (identified by Sequence ID No.):

1 and 5, 6 and 5, 7 and 5, 81 and 84,

85 and 88, 89 and 92 and 93 and 96

8. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising:

a composition according to claim 6 or 7, and

further comprising a polymerase.

9. A kit according to claim 8 wherein said polymerase is thermostable

10. A method for determining the presence of any human papilloma virus in a test sample, comprising:

a hybridizing DNA in the test sample with at least one consensus oligonucleotide selected from the group of sequences consisting of:

SEQ ID No.

1	CAGATGCTC	TGTGGCGGCC	TAGTG.
5	AGGTGTCAGG	AAAACCAAT	TTATT,
6	GAATTAGTTA	GACCATTTAA	AAG and
7	GGGAAACAC	CAGAATGGAT	A;

and their complements,

said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b. determining the presence of human papilloma virus by detecting the signal generated.

11. A method for determining the presence of human papilloma virus type 16 in a test sample, comprising

a. hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of sequences consisting of:

SEQ ID No.

81	GCTGCAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA,
93	GTATGGAACA	ACATTAGAAC	ASCA,
94	TGCTGTTCTA	ATGTTGTTCC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT and
96	AAATCACACA	ACGGTTTGTG	GTAT;

and their complements, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b. determining the presence of human papilloma virus by detecting the signal generated.

12. A method for determining the presence of human papilloma virus type 18 in a test sample, comprising.

a. hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of sequences consisting of

SEQ ID No.

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTICTAT	GTCTTGCACT	GAA,
87	CCTGTGTA?A	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG,
89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGCTTGCAA	TTTA,
91	TTACAGAGGT	ATTGGAATTT	GCAIT and
92	AATGCAAAIT	CAAATACCTC	TGTAA;

and their complements,

said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b. determining the presence of human papilloma virus by detecting the signal generated.

13. A method according to any of claims 10-12, further comprising a step of amplification prior to or concurrent with said hybridizing step

14. A method according to claim 13, wherein said amplification step comprises PCR or LCR.

Patentansprüche

Patentansprüche für folgende Vertragsstaaten: AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, NL, SE

1. Zusammensetzung, die für die LCR ("ligase chain reaction", Ligasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht:

LCR5: SEQ ID N°

81	GCTGCAAAACA	ACTATACATG	ATATAA.
82	TTATATCATG	TATAGTTGTT	TGCAGC.
83	TATTAGAAATG	TGTGTACTGC	AAGCA.
84	TGCTTGCAAT	ACACACATTG	TAATA.

LCR6: SEQ ID N°

85	CTTCACTGCA	AGACATAGAA	ATAA.
86	TTATTTCTAT	GTCTTGCAAT	GAA.
87	CCTGTGTATA	TTGCAAGACA	GTAT.
88	TACTGTCTTG	CAATATACAC	AGG.

LCR7: SEQ ID N°

89	TATATTGCAA	GACAGTATTG	GAAC.
90	GTTCGAATAC	TGCTTGCAA	TTTA.
91	TTACAGAGGT	ATTTGAATTT	GCATT.
92	AATGCAAAAT	CAAAATACCTC	TGTAA.

 und

LCR8: SEQ ID N°

93	GTATGGAACA	ACATTAGAAC	AGCA.
94	TGCTGTTCTA	ATGTTGTTCC	ATAC.
95	ATACAAACAA	CCGTTGTGTG	ATTT.
96	AAATCACACA	ACGGTTTGTT	GTAT.

2. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus Typ 16, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR5 (SEQ ID N° 81, 82, 83 und 84) und LCR8 (SEQ ID N° 93, 94, 95 und 96)
3. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus Typ 18, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR6 (SEQ ID N° 85, 86, 87 und 88) und LCR7 (SEQ ID N° 89, 90, 91 und 92).
4. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:
eine Zusammensetzung nach einem der Ansprüche 1 bis 3, und des weiteren eine Ligase.
5. Kit nach Anspruch 4, worin die Ligase thermostabil ist.
6. Zusammensetzung, die bei der PCR ("polymerase chain reaction" Polymerasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung folgendes umfaßt:

EP 0 477 972 B1

einen ersten Nukleinsäureprimer, der zur Richtung gleichläufig ist, welcher zur Hybridisierung an den gegenläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

5
1 CAGATGTC TC TGTGGCGGCC TAGTG.

6 GAATTAGTTA GACCATTTAA AAG,
10 7 GGGGAAACAC CAGAATGGAT A,
81 GCTGCAAAACA ACTATACATG ATATAA,
85 CTTCACCTGCA AGACATAGAA ATAA.
89 TATATTGCAA GACAGTATTG GAAC und
15 93 GTATGGAACA ACATTAGAAC AGCA; und

einen zweiten Nukleinsäureprimer, der zur Richtung gegenläufig ist, welcher zur Hybridisierung an den gleichläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

SEO ID Nr

25 5 AGGTGTCAGG AAAACCAAAT TTATT.
84 TGCTTGCAGT ACACACATTC TAATA,
88 TACTGTCCTG CAATATACAC AGG,
92 AATGCAAAAT CAAATACCTC TGTA und
30 96 AAATCACACA ACGGTTTGTT GTAT;

vorausgesetzt, daß der erste und der zweite Primer an ihre jeweiligen gleich- und gegenläufigen Stränge an solchen Stellen hybridisieren, daß ihre 3'-Enden nicht überlappen, und daß die 5'-Enden der Primer in Verlängerungsrichtung weiter räumlich abgesetzt sind als die 3'-Enden der Primer.

- 35 7. Zusammensetzung nach Anspruch 6, worin der erste und zweite Primer aus den folgenden Paaren von Oligonukleotidsequenzen (die durch die Sequenz ID Nr bezeichnet sind) gewählt sind:
1 und 5, 6 und 5, 7 und 5, 81 und 84,
85 und 88, 89 und 92, und 93 und 96.
- 40 8. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:
eine Zusammensetzung nach Anspruch 6 oder 7 und des weiteren eine Polymerase.
- 45 9. Kit nach Anspruch 8, worin die Polymerase thermostabil ist
10. Consensus-Oligonukleotid zur Hybridisierung der humanen papillomaviren Typ 6, 11, 16, 18, 31, 33 und 61, wobei das Oligonukleotid ungefähr 10 bis ungefähr 60 Oligonukleotide lang ist und aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht

SEO ID Nr

50
1 CAGATGTC TC TGTGGCGGCC TAGTG.
5 AGGTGTCAGG AAAACCAAAT TTATT,
55 6 GAATTAGTTA GACCATTTAA AAG und
7 GGGGAAACAC CAGAATGGAT A;

und aus deren Komplementen.

11. Typ-spezifisches Oligonukleotid zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 16, das eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus folgendem besteht:

SEQ ID Nr

```

81   GCTGCAAACA ACTATACATG ATATAA.
82   TTATATCATG TATAGTTGTT TGACGC.
83   TATTAGAATG TGTGTACTGC AAGCA.
84   TGCTTGCACT ACACACATTC TAATA.
93   GTATGGAACA ACATTAGAAC AGCA.
94   TGCTGTTCTA ATGTTGTTCC ATAC.
95   ATACAACAAA CCGTTGTGTG ATTT
96   AAATCACACA ACGGTTTGTT GTAT; und

```

und aus deren Komplementen.

12. Typ-spezifisches Oligonukleotid zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 18, das eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus folgendem besteht:

SEQ ID Nr

```

85   CTTCACTGCA AGACATAGAA ATAA.
86   TTATTTCAT  GTCCTGCAGT GAA.
87   CCTGTGTATA TTGCAAGACA GTAT.
88   TACTGTCTTG CAATATACAC AGG.
89   TATATTGCAA GACAGTATTG GAAC.
90   GTTCCAATAC TGTCTTGCAA TTAA.
91   TTACAGAGGT ATTTGAATTT GCATT
92   AATGCAAATT  CAAATACCTC TGTA; und

```

und aus deren Komplementen.

13. Verfahren zur Bestimmung der Anwesenheit irgendeines humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:

- a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Consensus-Oligonukleotid, das aus der Gruppe nach Anspruch 10 gewählt ist, wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und
b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

14. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 16 in einer Probe, das folgendes umfaßt:

- a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe nach Anspruch 11 gewählt ist, wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und
b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

15. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 18 in einer Testprobe, das folgendes umfaßt:

- a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe nach Anspruch 12 gewählt ist, wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und

b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

16. Verfahren nach einem der Ansprüche 13-15, das des weiteren einen Vervielfachungsschritt umfaßt, der vor oder in Konkurrenz mit dem Hybridisierungsschritt stattfindet

17. Verfahren nach Anspruch 16, worin der Vervielfachungsschritt PCR oder LCR umfaßt.

Patentansprüche für folgenden Vertragsstaat: ES

1. Zusammensetzung, die für die LCR ("ligase chain reaction", Ligasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht.

LCR5: SEQ ID N°

```

81   GCTGCAAAACA ACTATACATG ATATAA,
82   TTATATCATG TATAGTTGTT TGCASC,
83   TATTAGAAATG TGTGTACTGC AAGCA,
84   TGCTTGCACT ACACACATTC TAATA;
    
```

LCR6: SEQ ID N°

```

85   CTTCACCTGCA AGACATAGAA ATAA,
86   TTATTTCTAT GTCTTGCACT GAA,
87   CCTGTGTATA TTGCAAGACA GTAT,
88   TACTGTCTTG CAATATACAC AGG;
    
```

LCR7: SEQ ID N°

```

89   TATATTGCAA GACAGTATTG GAAC,
90   GTTCCAATAC TGTCTTGCAA TTTA,
91   TTACAGAGGT ATTGAATTT GCATT,
92   AATGCAAAAT CAAATACCTC TGTAA;
    
```

LCR8: SEQ ID N°

```

93   GTATGGAACA ACATTAGAAC AGCA, und
94   TGTGTTTCTA ATGTTGTTC ATAC,
95   ATACAACAAA CCGTTGTGTG ATTT,
96   AAATCACACA ACGGTTTGTT GTAT.
    
```

2. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus Typ 16 der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht. LCR5 (SEQ ID N° 81, 82, 83 und 84) und LCR6 (SEQ ID N° 85, 86, 87 und 88).

3. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus TYP 18 der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht. LCR6 (SEQ ID N° 85, 86, 87 und 88) und LCR7 (SEQ ID N° 89, 90, 91 und 92).

4. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:
eine Zusammensetzung nach einem der Ansprüche 1 bis 3, und des weiteren eine Ligase.

5. Kit nach Anspruch 4, worin die Ligase thermostabil ist.

6. Zusammensetzung, die bei der PCR ("polymerase chain reaction" polymerasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung folgendes umfaßt:

einen ersten Nukleinsäureprimer, der zur Fichtung gleichläufig ist, welcher zur Hybridisierung an den gegenläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

SEQ ID Nr

1 CAGATGTCCTC TGTGGCGGCC TAGTG.

6 GAATTAGTITA GACCATTTAA AAG.

7 GGGGAAACAC CAGAATGGAT A.

81 GCTGCAAAACA ACTATACATG ATATAA.

85 CTTCACCTGCA AGACATAGAA ATAA.

89 TATATTGCAA GACAGTATTG GAAC und

93 GTATGGAACA ACATTAGAAC AGCA; und

einen zweiten Nukleinsäureprimer, der zur Richtung gegenläufig ist, welcher zur Hybridisierung an den gleichläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

SEQ ID Nr

5 AGGTGTCAGG AAAACCAAAAT TTATT.

84 TGCTTGCACT ACACACATTC TAATA.

88 TACTGTCTTG CAATATACAC AGG.

92 AATGCAAAAT CAAATACCTC TGTA und

96 AAATCACACA ACGGTTTGTT GTAT;

vorausgesetzt, daß der erste und der zweite Primer an ihre jeweiligen gleich- und gegenläufigen Stränge an solchen Stellen hybridisieren, daß ihre 3'-Enden nicht überlappen, und daß die 5'-Enden der Primer in Verlängerungsrichtung weiter räumlich abgesetzt sind als die 3'-Enden der Primer

7. Zusammensetzung nach Anspruch 6, worin der erste und zweite Primer aus den folgenden Paaren von Oligonukleotidsequenzen (die durch die Sequenz ID Nr bezeichnet sind) gewählt sind:

1 und 5, 6 und 5, 7 und 5, 81 und 84,

85 und 88, 89 und 92, und 93 und 96

8. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:

eine Zusammensetzung nach Anspruch 6 oder 7, und des weiteren eine Polymerase.

9. Kit nach Anspruch 8, worin die Polymerase thermostabil ist.

10. Verfahren zur Bestimmung der Anwesenheit irgendeines humanen papillomavirus in einer Testprobe, das folgendes umfaßt:

a Hybridisieren der DNA in der Testprobe mit wenigstens einem Consensus-Oligonukleotid, das aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht:

SEQ ID Nr

5

10

```

1   CAGATGTC TC TGTGGCGGCC TAGTG.
5   AGGTGTCAGG AAAACCAAAT TTATT
6   CAATTAGTA GACCAATTAA AAG und
7   GGGGAAACAC CAGAATGGAT A;

```

15

und aus deren Komplementen,

wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und

b Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

20

11. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 16 in einer Probe, das folgendes umfaßt

a Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht.

25

SEQ ID Nr

30

35

40

```

81   GCTGCAACA ACTATACATG ATATAA,
82   TTATATCATG TATAGTTGTT TGCAGC,
83   TATTAGAATG TGTGTACTGC AAGCA,
84   TGCTTGCACT ACACACATTC TAATA,
93   GTATGGAACA ACATTAGAAC AGCA,
94   TGCTGTTCTA ATGTTGTTCC ATAC,
95   ATACAACAAA CCGTTGTGTG ATT und
96   AAATCACACA ACGGTTTGTT GTAT;

```

und aus deren Komplementen,

wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und

b Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

50

12. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 18 in einer Testprobe, das folgendes umfaßt

a Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht

55

SEQ ID Nr

5 85 CTTCACTGCA AGACATAGAA ATAA.
 86 TTATTTCTAT GTCTTGCAGT GAA,
 87 CCTGTGTATA TTGCAAGACA GTAT.
 88 TACTGTCTTG CAATATACAC AGG.
 89 TATATTGCAA GACAGTATTG GAAC.
 10 90 GTTCCAATAC TGTCTTGCAA TTTA,
 91 TTACAGAGGT ATTTGAATTT GCATT und
 92 AATGCAAAAT CAAATACCTC TGTAAT;

15 und aus deren Komplementen,
 wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nach-
 weisbaren Signals befähigt ist, und
 b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

20 13. Verfahren nach einem der Ansprüche 10-12, das des weiteren einen Vervielfachungsschritt umfaßt, der vor oder
 in Konkurrenz mit dem Hybridisierungsschritt stattfindet.

14. Verfahren nach Anspruch 13, worin der vervielfachungsschritt PCR oder LCR umfaßt

25 **Revendications**

Revendications pour les Etats contractants suivants : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, NL, SE

30 1. Composition utile dans la LCR pour amplifier l'ADN de virus du papillome humain présent dans échantillon à doser,
 ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes
 étant sélectionnées dans le groupe constitué par les ensembles d'oligonucléotides suivants :

35 LCR5 : n° d'identification
 81 GCTGCAAACA ACTATACATG ATATAA,
 82 TTATATCATG TATAGTTGTT TGCAGC,
 83 TATTAGAATG TGTGTACTGC AAGCA,
 40 84 TGCTTGCAGT ACACACATTC TAATA ;

45 LCR6 : n° d'identification
 85 CTTCACTGCA AGACATAGAA ATAA,
 86 TTATTTCTAT GTCTTGCAGT GAA,
 87 CCTGTGTATA TTGCAAGACA GTAT,
 50 88 TACTGTCTTG CAATATACAC AGG ;

55

LCR7 : n° d'identification

89 TATATTGCAA GACAGTATTG GAAC
 90 GTTCCAATAC TGTCTTGCAA TTTA,
 91 TTACAGAGGT ATTTGAAATTT GCATT,
 92 AATGCAAATT CAAATACCTC TGTA ; et

LCR8 : n° d'identification

93 GTATGGAACA ACATTAGAAC AGCA,
 94 TGCTGTTCTA ATGTTGTTC ATAC,
 95 ATACAACAAA CCGTTGTGTG ATTT,
 96 AAATCACACA ACGGTTTGT GTAT.

2. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 16 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants

LCR5 (n° d'identification 81, 82, 83 et 84) et LCR8 (n° d'identification 93, 94, 95 et 96)

3. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 18 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants

LCR6 (n° d'identification 85, 86, 87 et 88) et LCR7 (n° d'identification 89, 90, 91 et 92).

4. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant une composition selon l'une quelconque des revendications 1 à 3, et en outre une ligase

5. Kit selon la revendication 4, dans lequel ladite ligase est thermostable

6. Composition utile dans la PCR pour amplifier l'ADN de virus du papillome humain présent dans un échantillon à doser, ladite composition comprenant :

une première amorce d'acide nucléique de direction sens, capable de s'hybrider au brin antisens de l'ADN de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

N° d'identification

1 CAGATGTCTC TGTGGCGGCC TAGTG,
 6 GAATTAGTTA GACCATTTAA AAG,
 7 GGGGAAACAC CAGAATGGAT A,
 81 GCTGCAAACA ACTATACATG ATATAA,
 85 CTTCACTGCA AGACATAGAA ATAA,
 89 TATATTGCAA GACAGTATTG GAAC et
 93 GTATGGAACA ACATTAGAAC AGCA ; et

une deuxième amorce d'acide nucléique de direction antisens, capable de s'hybrider au brin sens de l'ADN

de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

5	N° d'identification 5	AGGTGTCAGG AAAACCAAAT TTATT,
	84	TGCTTGCACT ACACACATTC TAATA,
	88	TACTGTCTTG CAATATACAC AGG,
10	92	AATGCAAATT CAAATACCTC TGTA et
	96	AAATCACACA ACGGTTTGT GTAT ;

pour autant que lesdites première et deuxième amorces s'hybrident à leurs brins respectifs antisens et sens à des emplacements tels que leurs extrémités 3' ne se chevauchent pas et que, dans la direction d'extension, les extrémités 5' desdites amorces soient plus espacées que les extrémités 3' desdites amorces.

- 15 7. Composition selon la revendication 5, dans laquelle lesdites première et deuxième amorces sont sélectionnées parmi les paires suivantes de séquences oligonucléotidiques (identifiées par leur numéro d'identification) :
- 20 1 et 5, 6 et 5, 7 et 5, 81 et 84, 85 et 88, 89 et 92, et 93 et 96
8. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant une composition selon la revendication 6 ou 7 et en outre une polymérase.
- 25 9. Kit selon la revendication 8, dans lequel ladite polymérase est thermostable.
10. Oligonucléotide consensus pour hybridation du virus du papillome humain des types 6, 11, 16, 18, 31, 33 et 61, lequel oligonucléotide a d'environ 10 à environ 60 nucléotides de long et est sélectionné dans le groupe de séquences constitué par :

	N° d'identification	
	1	CAGATGCTC TGTGGCGGCC TAGTG,
	5	AGGTGTCAGG AAAACCAAAT TTATT,
35	6	GAATTAGTTA GACCATTTAA AAG et
	7	GGGGAACAC CAGAATGGAT A ;

et leurs compléments

- 40 11. Oligonucléotide spécifique d'un type, destiné à déterminer la présence du virus du papillome humain de type 16, ayant une séquence sélectionnée dans le groupe constitué par

45	N° d'identification	
	81	GCTGCAACA ACTATACATG ATATAA,
	82	TTATATCATG TATAGTTGTT TGCAGC,
	83	TATTAGAAATG TGTGTACTGC AAGCA,
	84	TGCTTGCACT ACACACATTC TAATA,
	93	GTATGGAACA ACATTAGAAC AGCA,
50	94	TGCTGTCTTA ATGTTGTTC ATAC,
	95	ATACAACAAA CCGTTGTGTG ATTT et
	96	AAATCACACA ACGGTTTGT GTAT ;

et leurs compléments

- 55 12. Oligonucléotide spécifique d'un type, destiné à déterminer la présence du virus du papillome humain de type 18, ayant une séquence sélectionnée dans le groupe constitué par :

N° d'identification

	85	CTTCACTGCA	AGACATAGAA	ATAA,
5	86	TTATTICTAT	GTCTTGCACT	GAA,
	87	CCTGTGTATA	TTGCAAGACA	GTAT,
	88	TACTGTCTTG	CAATATACAC	AGG,
	89	TATATTGCAA	GACAGTATTG	GAAC,
	90	GTTCCAATAC	TGTCITGCAA	TITA,
10	91	TTACAGAGGT	ATTITGAATTT	GCATT et
	92	AATGCAAAIT	CAAATACCTC	TGTAA ;

et leurs compléments

- 15 13. Procédé de détermination de la présence d'un virus quelconque du papillome humain dans un échantillon à doser, comprenant :
- 20 a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide consensus sélectionné dans le groupe selon la revendication 10, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et
- b. la détermination de la présence du virus du papillome humain par détection du signal émis
14. Procédé de détermination de la présence du virus du papillome humain de type 16 dans un échantillon à doser, comprenant :
- 25 a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe selon la revendication 11, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et
- 30 b. la détermination de la présence du virus du papillome humain par détection du signal émis.
15. Procédé de détermination de la présence du virus du papillome humain de type 18 dans un échantillon à doser, comprenant :
- 35 a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe selon la revendication 12, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et
- 40 b. la détermination de la présence du virus du papillome humain par détection du signal émis.
16. Procédé selon une quelconque des revendications 13 à 15, comprenant en outre une étape d'amplification avant ou pendant ladite étape d'hybridation.
17. Procédé selon la revendication 16, dans lequel ladite étape d'amplification comprend la PCR ou la LCR.

45 **Revendications pour l'Etat contractant suivant : ES**

- 50 1. Composition utile dans la LCR pour amplifier l'ADN de virus du papillome humain présent dans échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants

LCR5 : n° d'identification

	81	GCTGCAAACA	ACTATACATG	ATATAA,
	82	TTATATCATG	TATAGTTGTT	TGCAGC,
55	83	TATTAGAATG	TGTGTACTGC	AAGCA,
	84	TGCTTGCAGT	ACACACATTC	TAATA ;

LCR6 : n° d'identification

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GTCTTGCAGT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG ;

LCR7 : n° d'identification

89	TATATTGCAA	GACAGTATTG	GAAC
90	GTTCCAATAC	TGTCITGCAA	TTTA,
91	TTACAGAGGT	ATTTGAATTT	GCATT,
92	AATGCAAATT	CAAATACCTC	TGTAA ; et

LCR8 : n° d'identification

93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTCTA	ATGTTGTTC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT,
96	AAATCACACA	ACGGTTTGTG	GTAT.

2. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 16 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

LCR5 (n° d'identification 81, 82, 83 et 84) et LCR8 (n° d'identification 93, 94, 95 et 96).

3. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 18 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

LCR6 (n° d'identification 85, 86, 87 et 88) et LCR7 (n° d'identification 89, 90, 91 et 92).

4. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant une composition selon l'une quelconque des revendications 1 à 3, et en outre une ligase

5. Kit selon la revendication 4, dans lequel ladite ligase est thermostable.

6. Composition utile dans la PCR pour amplifier l'ADN de virus du papillome humain présent dans un échantillon à doser, ladite composition comprenant :

une première amorce d'acide nucléique de direction sens, capable de s'hybrider au brin antisens de l'ADN de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

N° d'identification

5	1	CAGATGTCCTC	TGTGGCGGCC	TAGTG,
	6	GAATTAGTTA	GACCATTTAA	AAG,
	7	GGGGAAACAC	CAGAATGGAT	A,
10	81	GCTGCAAAAC	ACTATACATG	ATATAA,
	85	CTTCACTGCA	AGACATAGAA	ATAA,
	89	TATATTGCAA	GACAGTATTG	GAAC et
	93	GTATGGAACA	ACATTAGAAC	AGCA ; et

15 une deuxième amorce d'acide nucléique de direction antisens, capable de s'hybrider au brin sens de l'ADN de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

N° d'identification

20	5	AGGTGTCAGG	AAAACCAAAT	TTATT,
	84	TGCTTGCACT	ACACACATTC	TAATA,
	88	TACTGTCTTG	CAATATACAC	AGG,
25	92	AATGCAAATT	CAAAATACCTC	TGTAA et
	96	AAATCACACA	ACGGTTTGTT	GTAT ;

pour autant que lesdites première et deuxième amorces s'hybrident à leurs brins respectifs antisens et sens à des emplacements tels que leurs extrémités 3' ne se chevauchent pas et que, dans la direction d'extension, les extrémités 5' desdites amorces soient plus espacées que les extrémités 3' desdites amorces

7. Composition selon la revendication 6, dans laquelle lesdites première et deuxième amorces sont sélectionnées parmi les paires suivantes de séquences oligonucléotidiques (identifiées par leur numéro d'identification) :

1 et 5, 6 et 5, 7 et 5, 81 et 84,

85 et 88, 89 et 92, et 93 et 96.

8. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant une composition selon la revendication 6 ou 7 et en outre une polymérase.

9. Kit selon la revendication 8, dans lequel ladite polymérase est thermostable

10. Procédé de détermination de la présence d'un virus quelconque du papillome humain dans un échantillon à doser comprenant :

a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide consensus sélectionné dans le groupe de séquences constitué par :

N° d'identification

50	1	CAGATGTCCTC	TGTGGCGGCC	TAGTG,
	5	AGGTGTCAGG	AAAACCAAAT	TTATT,
	6	GAATTAGTTA	GACCATTTAA	AAG et
	7	GGGGAAACAC	CAGAATGGAT	A ;

55 et leurs compléments, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et

b. la détermination de la présence du virus du papillome humain par détection du signal émis.

11. Procédé de détermination de la présence du virus du papillome humain de type 16 dans un échantillon à doser, comprenant :

a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe de séquences constitué par :

N° d'identification

81	GCTGCAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA,
93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT et
96	AAATCACACA	ACGGTTTGTT	GTAT ;

et leurs compléments,

ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et

b. la détermination de la présence du virus du papillome humain par détection du signal émis.

12. Procédé de détermination de la présence du virus du papillome humain de type 18 dans un échantillon à doser, comprenant :

a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe de séquences constitué par :

N° d'identification

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTCTAT	GTCTTGCACT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG,
89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGTCTTGCAA	TTTA,
91	TTACAGAGGT	ATTTGAATTT	GCATT et
92	AATGCAAATT	CAAATACCTC	TGTAA ;

et leurs compléments,

ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et

b. la détermination de la présence du virus du papillome humain par détection du signal émis.

13. Procédé selon une quelconque des revendications 10 à 12, comprenant en outre une étape d'amplification avant ou pendant ladite étape d'hybridation.

14. Procédé selon la revendication 13, dans lequel ladite étape d'amplification comprend la PCR ou la LCR.

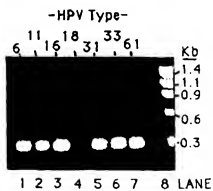


FIG. 1

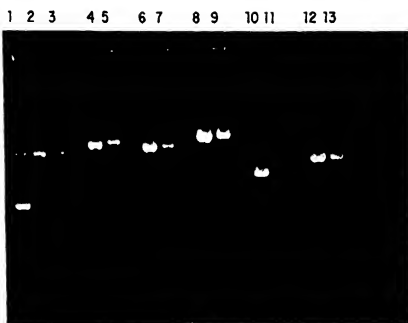


FIG. 2

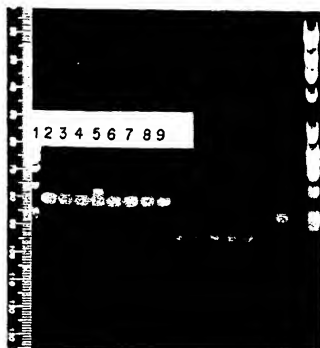


FIG. 3

FIG. 4

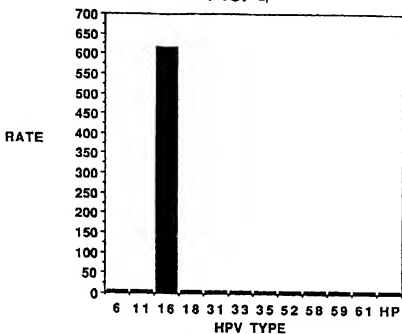


FIG. 5

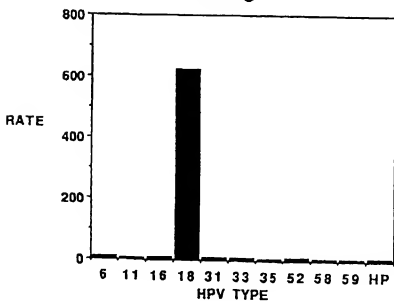


FIG. 6

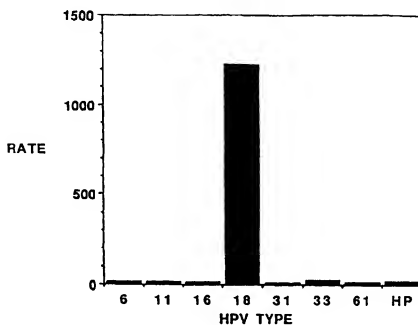


FIG. 7

